RECHERCHE EN COMPORTEMENT SUR TELIDON

3 Étude des caractéristiques humaines à l'utilisation de représentations graphiques sur Telidon

86-0278/

ÉTUDE DES RÉACTIONS HUMAINES À L'UTILISATION DE REPRÉSENTATIONS GRAPHIQUES SUR TÉLIDON

par
Michael I. Mills
UNIVERSITÉ DE MONTRÉAL

POUR
Le Groupe de recherche et d'évaluation en comportement
Direction de la technologie d'information
Ministère des Communications
Ottawa (Canada)

N° du contrat: 09SU.36001-0-3165

Des exemplaires de cette étude sont
disponibles, sans frais, auprès de la:

Direction de l'information
Ministère des Communications
300, rue Slater
Ottawa (Canada)
K1A 0C8

Numéro de rapport: DOC-TBR-DBRE-81-2-F

SOMMAIRE

L'emploi d'images et de graphiques sur Télidon, bien que riche de promesses, peut imposer aux fournisseurs d'information, aux exploitants et aux utilisateurs du système, des coûts qui peuvent influer sur l'acceptabilité du service. Ces coûts, qui peuvent varier selon le genre et la complexité de l'image, nous amènent à nous demander quels avantages découleraient de l'emploi d'images visuelles sur Télidon. Le présent rapport tente d'abord de faire le point des recherches qui s'effectuent sur les processus de cognition et de perception pour apprendre comment les hommes comprennent et utilisent les images visuelles; et ensuite, d'établir l'importance de ces travaux pour l'utilisation d'images et de graphiques sur Télidon. Chacune des trois grandes parties du rapport développe un certain nombre de thèmes qui, sans l'expliquer totalement, éclairent quelque peu la question de l'imagerie visuelle sur Télidon.

Le premier chapitre est consacré à l'étude des avantages relatifs des images et des mots dans le processus de communication. Une image peut-elle remplacer un énoncé verbal? Dès qu'on tente de répondre à cette question, on constate qu'il existe des différences fondamentales entre les images et les mots et qu'une image peut difficilement avoir les mêmes fonctions descriptives que le verbe. Le premier chapitre expose aussi quelques notions théoriques qui peuvent aider à expliquer pourquoi l'utilisation d'images pose des problèmes difficiles pour certaines applications de Télidon comme le classement et le recouvrement automatiques d'images par ordinateur et l'élaboration de répertoires iconographiques. En conclusion, ce chapitre soutient que le véritable défi posé aux fournisseurs d'information n'est pas de trouver des moyens de remplacer les mots par des images, mais plutôt de comprendre comment l'image peut servir de base conceptuelle pour faciliter la compréhension du texte et, réciproquement, comment le texte peut faciliter celle de l'image.

Le deuxième chapitre étudie la question suivante : faut-il posséder des compétences particulières pour "lire" une image Télidon? En d'autres mots, l'expérience directe du monde visuel fournit-elle les compétences nécessaires pour "lire" une image Télidon? Selon certains chercheurs, on pourrait percevoir plus facilement et plus rapidement les images s'éloignant du réalisme photographique (par exemple les dessins-contours, les bandes dessinées et les caricatures) que les photographies de la chose elle-même. Les recherches sur la perception des dessins-contours et des caricatures révèlent un certain nombre de compromis (par exemple, entre le délai de transmission d'une image et le réalisme de l'image) qui permettraient d'exploiter le point fort de Télidon en tant que médium pictural, c'est-à-dire sa capacité de produire des graphiques de grande qualité.

Le troisième chapitre porte sur l'utilisation de schémas graphiques pour communiquer des idées complexes. Il donne un aperçu du processus qu'entraîne l'utilisation de plans ou de cartes topographiques et soulève la question des différences individuelles quant à la compréhension de certains types d'images. Nous examinons aussi la raison pour laquelle des schémas graphiques simples comme les diagrammes de Venn peuvent aider à penser et à apprendre. Il semble que si les représentations graphiques abstraites facilitent l'apprentissage, c'est qu'elles permettent de représenter sous une forme visible des propriétés et des relations qui ne sont pas directement observables dans l'apparence matérielle des objets et des événements. Nous avançons, enfin, que des schémas graphiques <u>dynamiques</u> qui illustrent des transformations favorisent davantage la pensée constructive et la résolution de problèmes que des graphiques statiques. Il serait peut-être possible de représenter de telles transformations dynamiques sur Télidon, de façon économique, au moyen d'une série d'images fixes, sans créer l'illusion du mouvement continu.

L'annexe A présente une Introduction à la psychologie de la perception de l'image spécialement rédigée pour le présent rapport par Adam Gopnik.

Les opinions et les interprétations que contient ce document n'engagent que l'auteur et ne reflètent pas nécessairement celles du ministère des Communications.

REMERCIEMENTS

Je voudrais remercier Madame Dorothy Phillips du ministère des Communications. À titre de responsable scientifique, Mme Phillips a contribué, par ses conseils, à préciser les idées exprimées dans ce rapport. Je tiens également à remercier d'autres membres du Groupe de recherche en comportement, messieurs Eric Lee, Paul Muter, Tom Whalen et Bill Treurniet, qui m'ont fourni des renseignements précieux au cours d'entretiens. Je veux remercier tout particulièrement M. Adam Gopnik, mon chargé de recherche qui a assuré la compilation de la documentation, et participé pleinement à la formulation des idées. Il est l'auteur de l'Introduction à la psychologie de la perception de l'image qui est présentée à l'annexe A.

TABLE DES MATIÈRES

		Page
SOMMAIRE		iii
REMERCIEMENTS		vii
INTRODUCTION		
0.1.	Pourquoi étudier le rôle des images visuelles sur Télidon?	1
0.1.1.	Produire des images à l'aide de techniques élémentaires de dessin	2
0.2.	Trois thèmes	7
CHAPITRE I - UNE IMAGE SUR TÉLIDON VAUDRA-T-ELLE MILLE MOTS?		
1.0.	Un argument courant concernant les images et les mots	13
1.1.	Une image peut-elle remplacer un énoncé verbal?	15
1.1.1.	Une image vaut mille mots, mais lesquels?	18
1.2.	Polysémie des images: Les "schèmes mentaux" dans la perception	20
1.2.1.	Schémas et transformations	21
1.3.	Descriptions d'images destinées aux machines	27
1.4.	Les images peuvent-elles servir d'index?	30
1.4.1.	Icônes, signes et symboles	32
1.5.	L'importance du contexte	39
1.5.1.	Qu'apportent les mots aux images?	40
1.5.2.	Qu'apportent les images aux mots?	43
1.5.3.	Interaction des mots et des images	46
1.6.	Résumé et conclusions	51
CHAPITRE II - LIRE LES IMAGES: LE RÉALISME EST-IL CE QU'IL Y A DE MIEUX?		
2.0.	Apprenons-nous à lire les images?	55
2.1.	Le langage des lignes	61
2.2.	L'équilibre temps/réalisme: les répercussions de l'utilisation de dessins-contours sur Télidon	66
2.3.	Les photographies, les dessins, les bandes dessinées, les caricatures	74
2.3.1.	Les bandes dessinées concrétisent les schèmes intérieurs	77
2.4.	Résumé et conclusions	88
CHAPITRE III - L'IMAGE AU SERVICE DE LA PENSÉE: DES SCHÈMES GRAPHIQUES POUR VÉHICULER DES IDÉES COMPLEXES		
3.0.	Comment les cartes nous renseignent	92
3.0.1.	Aptitude à lire les cartes: différences entre les individus	96
3.0.2.	Cartes cinématographiques	98
3.1.	Schèmes graphiques et pensée	102
3.1.2.	Graphisme dynamique	107
3.1.3.	Mouvement et séquence	113
3.2.	Résumé et conclusions	114

Page

RÉSUMÉ ET CONCLUSIONS..... 119
OUVRAGES DE RÉFÉRENCE..... 125

ANNEXE A: INTRODUCTION À LA PSYCHOLOGIE DE LA PERCEPTION DE L'IMAGE..... 131

ANNEXE B: BIBLIOGRAPHIE SUPPLÉMENTAIRE..... 151

TABLE DES FIGURES

Page

Figure:

0.1 - Les parties de l'image apparaissent sur l'écran de visualisation de l'utilisateur dans l'ordre indiqué..... 5

1.0 - Le chat est sur la carpette..... 17

1.1 - Analogie d'une bande dessinée..... 25

1.2 - Dessin de Steinberg..... 28

1.3 - Pictogrammes: Jeux olympiques de Montréal..... 37

1.4 - "Droodles" (dessins-attrapes): a) Un nain jouant du trombone dans une cabine téléphonique; b) Une autruche qui s'est aperçue trop tard que c'était du sable mouvant..... 41

1.5 - Dessin extrait de l'expérience de Bransford et Johnson (1972)..... 45

1.6 - Exemples des quatre types de dessins qui accompagnaient les paires de mots présentées dans l'expérience de Wollen et alii. (D'après Wollen, Weber et Lowry, 1972)..... 47

1.7 - Un seul dessin qui représente deux événements distincts: (a) Un pied écrasant un insecte (vue de profil); (b) Une voiture entrant en collision avec un poteau (vue en plongée)..... 49

2.0 - Le dessin (a) peut représenter un cube: les lignes convergentes respectent les lois de la perspective linéaire. Le dessin (b) ne peut pas, à proprement parler, représenter un cube, car les lignes qui en prolongent les côtés sont parallèles..... 58

2.1 - La signification d'une ligne..... 63

2.2 - Dessin-contour d'une maison fait à partir d'une photographie en couleur..... 69

2.3 - Une matrice d'images destinée à un survol de l'information..... 73

Page

2.4 - Trois des quatre types d'images utilisées dans l'expérience de Ryan et Schwartz..... 75

2.5 - Deux projections possibles d'un cube..... 79

2.6 - Calques de photographies et caricatures de Richard Nixon..... 83

2.7 - Exemples des photographies et des caricatures exployées dans l'expérience de Perkins et de Hagen (1980)..... 86

3.0 - Plan de métro..... 94

3.1 - Disposition réelle des pièces d'un appartement et disposition de ces pièces reproduite de mémoire par une fillette de douze ans..... 95

3.2 - Métaphore visuelle du concept "mariage réussi/ mariage raté"..... 105

3.3 - Solution d'un problème géométrique..... 109

3.4 - Images fixes tirées d'un film sur les "symboles dynamiques"..... 111

INTRODUCTION

0.1 Pourquoi étudier le rôle des images visuelles sur Télidon?

Télidon est un service de recherche documentaire mis au point par le ministère des Communications du Canada. Il permet à l'utilisateur de consulter, au moyen d'un téléviseur ordinaire et d'un mini-clavier, des banques de données stockées dans un ordinateur central. Vidéotex[1] est le terme général qui désigne un service interactif de recherche documentaire. Le vidéotex offre une vaste gamme de possibilités dans le domaine de l'information: il permet à l'utilisateur de se renseigner sur des sujets d'intérêt général comme l'actualité, la météo et les cours de la bourse. On peut aussi envisager des interactions plus spécialisées comme l'achat d'une maison, les jeux vidéo, l'établissement d'une déclaration des revenus, la préparation d'un repas et même des leçons de mathématiques.

Télidon n'est pas seulement un système d'information écrite. Il peut aussi fournir des renseignements sous forme d'images et de graphiques. Toutes sortes d'images sont accessibles à la télévision, sur les panneaux-réclame, dans les films, les revues, les livres et ailleurs. Elles servent à de nombreuses fins, tant pédagogiques que publicitaires. L'image visuelle sur Télidon est différente et son rôle comporte quelques éléments problématiques. Afin de comprendre pourquoi il en est ainsi, nous examinerons, dans un premier temps, l'élaboration des images et des textes sur Télidon.

(Une description sommaire de la méthode de création de pages suit. Bown et al. (1979) ont déjà décrit de façon précise le processus de création, de mémorisation et d'extraction des images, des textes et des représentations graphiques.)

---

[1] Voir l'ouvrage de Woolfe (1980) qui donne une description détaillée de la nature du vidéotex, comparant les caractéristiques de Télidon à celles des systèmes vidéotex européens et américains. Les publications de Bown, O'Brien, Sawchuck et Storey (1978, 1979) décrivent le terminal Télidon ainsi que ses caractéristiques de transmission.

0.1.1. Produire des images à l'aide de techniques élémentaires de dessin

Imaginez que vous êtes, outre un consommateur, un fournisseur d'information désireux de créer quelques pages d'information sur un sujet donné, par exemple la navigation à voile. Vous devrez d'abord résoudre une série de problèmes sur la façon de concevoir des pages d'information pour permettre à l'utilisateur de saisir rapidement les points essentiels de votre message. Votre premier problème pourrait être la conception d'une structure de présentation des données qui rende bien les relations conceptuelles entre ces points essentiels. Ensuite, vous devrez choisir un moyen de communication pour les présenter à l'utilisateur. Pour le renseigner sur les bateaux à voile, quels aspects de l'information faudrait-il présenter au moyen de textes, au moyen de graphiques ou au moyen d'une combinaison quelconque de textes et de graphiques? La création d'une page de texte décrivant comment faire de la voile ne présente aucune difficulté particulière. Vous devez évidemment décider de ce qu'il faut dire et de la meilleure façon de le dire, ce qui pourrait exiger un peu d'imagination. Cela fait, vous n'avez qu'à taper le texte à l'aide du terminal de création de pages, vos décisions se limitant à la mise en page, à la couleur et à d'autres détails. La production de caractères alphanumériques sur Télidon dépend d'un procédé relativement simple.

Supposons maintenant que vous voulez ajouter un dessin simple aux pages documentaires (voir la figure 1). Vous devrez alors recourir à un langage graphique particulier nommé Instructions de description de l'image (IDI). Les IDI sont une série de commandes permettant de préciser, à l'aide d'un photostylet ou des touches d'un clavier, les dimensions et les formes des différentes parties d'une image ainsi que l'endroit où elles apparaîtront sur l'écran. Ils peuvent servir à dessiner des figures géométriques primitives comme le point, la ligne,

l'arc, le polygone, et à choisir des tons et des couleurs. Ainsi, pour dessiner la petite marine de la figure 0.1, vous pourriez commencer par définir sur l'écran à dessin, à l'aide de l'instruction POLYGONE, les sommets de la coque du voilier; l'ordinateur relie ensuite les points des sommets pour faire apparaître la forme de la coque. Pour dessiner le soleil, vous utiliseriez ensuite la commande ARC pour définir le point de départ, un point central et un point d'arrêt; une fois interprétée par l'ordinateur, cette commande ferait jaillir un "soleil" circulaire sur l'écran. Les autres parties de l'image comme le mât et la ligne d'horizon seraient créés de la même façon, par le choix et l'exécution de commandes IDI. L'ordre dans lequel nous venons de décrire les parties de l'image (figure 0.1) sera celui de leur apparition sur l'écran de l'utilisateur au moment de l'affichage du dessin. Or, si vous aviez dessiné les parties de l'image dans l'ordre indiqué par les chiffres de la figure 0.1 (d'abord la coque, puis le soleil et ensuite la ligne d'horizon) c'est dans cet ordre que les segments seraient mémorisés dans la description informatisée de l'image et que l'utilisateur les verrait apparaître sur son écran: c'est-à-dire successivement. Quant à la figure 0.1, l'apparition de la séquence serait rapide; une affaire de secondes tout au plus. En ce qui concerne les images complexes, il faudrait utiliser un plus grand nombre d'IDI pour décrire l'image et il faudrait plus de temps pour que l'image se dessine sur l'écran du terminal de l'utilisateur.

On remarquera que pour emmagasiner et afficher des images photographiques sur Télidon, on peut passer directement au mode alpha-photographique ou de visualisation par profil de multiplets. Ce mode permet d'encoder une photographie au moyen d'une multitude de points individuels ou éléments imagiers, au lieu de recourir aux figures géométriques primitives qui servent à dessiner une image graphique. Au moment d'extraire l'image, ces points ou éléments imagiers seraient

reportés directement sur l'écran de télévision de l'utilisateur par balayage lent ou en mode de transmission de fac-similé; dans ce cas, l'image remplit graduellement l'écran ligne par ligne de haut en bas ou de bas en haut. Cette méthode de visualisation par profil de multiplets permet la transmission d'images d'une qualité photographique sans utiliser les Instructions de description de l'image. Toutefois, comme le montreront les chapitres subséquents, elle comporte des délais de transmission (l'affichage d'une image photographique selon le mode de représentation par profil de multiplets* exige plusieurs minutes) et d'autres exigences de stockage.

L'emploi des IDI pour décrire les images a de grands avantages par rapport aux autres méthodes d'exécution de graphiques par exemple celles de certains systèmes européens. Ainsi les IDI permettent de tracer des graphiques de qualité supérieure et de décrire des images sous une forme qui pourra s'adapter à d'éventuelles modifications des caractéristiques des terminaux. Différents terminaux ayant une capacité plus ou moins grande de résolution pourraient interpréter les IDI une fois munis du décodeur approprié (voir Bown et alii, 1978). Cependant, ce qui importe aux fins du présent rapport, c'est de comprendre que la fourniture d'information sous forme d'images visuelles, que ce soit au moyen d'IDI ou en mode de visualisation par profil de multiplets, n'est pas semblable à la façon normale de dessiner des images sur du papier ou de photographier des images pour la télévision. Il y a lieu d'envisager

* Le délai de l'apparition d'une image est fonction de la vitesse de transmission. Des lignes de transmission à vitesse supérieure à celle des lignes dont les usagers disposent actuellement élimineraient ce problème.

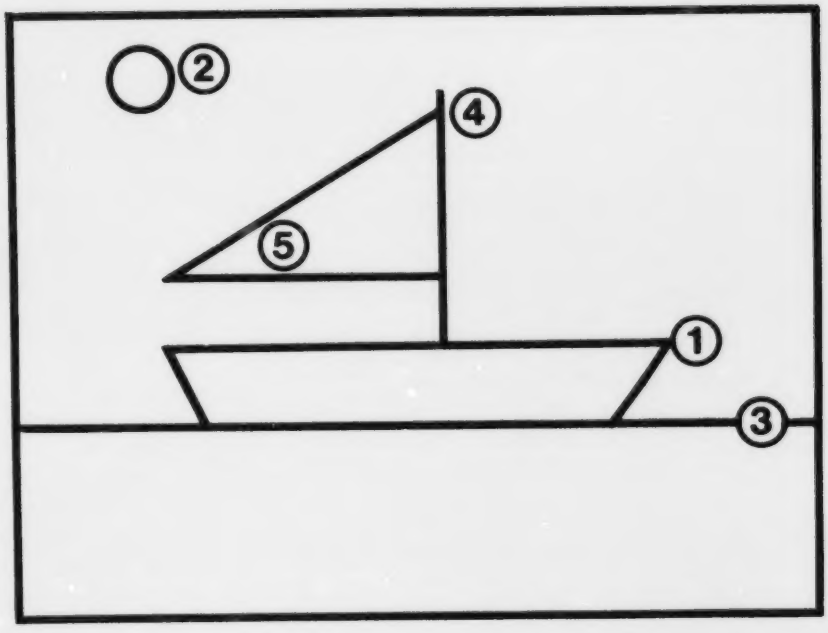


Figure 0.1 Les éléments de l'image apparaissent sur l'écran de l'utilisateur dans l'ordre indiqué.

certains frais et certains facteurs de pondération qu'on pourrait résumer comme suit:

1) Un fournisseur d'information qui désire concrétiser une idée dans une image visuelle plutôt que dans un texte peut encourir des frais en termes de temps et d'efforts consacrés à la conception et à la construction d'une image au moyen des IDI. En général, dans le cas de Télidon, la création d'une image semble plus compliquée que l'introduction d'un texte, le temps et l'effort augmentant en fonction de la complexité de l'image.

2) Même si une image visuelle peut fournir des renseignements précieux à un utilisateur, elle peut aussi entraîner des frais reliés à la réception de cette image. Il faut tenir compte du temps que l'utilisateur passe à attendre que les IDI construisent graduellement l'image sur l'écran. Plus l'image est chargée, plus l'attente est longue. En général, la transmission d'images graphiques simples définies au moyen d'IDI prend beaucoup moins de temps (secondes) que celle d'images photographiques transmises par profil de multiplets (minutes). On peut aussi compter les efforts cognitifs qu'exigera la compréhension d'une image ou d'un graphique mal conçu ou, inversement, la perte de temps que représente une image complexe qui, loin de favoriser la compréhension du message, n'est rien de plus que du tape-à-l'oeil. L'emploi d'images sur Télidon peut souvent entraîner des frais de ce genre.

3) Enfin, l'emploi d'images peut être coûteux pour l'exploitant du système ou le fournisseur d'information puisqu'il faut de la mémoire pour entreposer une description d'image dans l'ordinateur central. Ces frais de stockage augmentent avec

la complexité du graphique, mais les images graphiques décrites sous la forme d'IDI demandent moins d'espace que la multitude d'éléments imagiers nécessaires pour décrire des graphiques ou des photographies en mode de disposition séquentielle.

0.2. Trois thèmes

Ces généralisations aident à comprendre pourquoi on peut vouloir étudier le rôle de la représentation graphique à l'aide de Télidon. En dépit des perspectives très intéressantes qu'offre l'utilisation d'images et de graphiques, les exploitants du Télidon, les fournisseurs d'information et les utilisateurs devront y dépenser des sommes qui pourraient bien affecter l'acceptation globale du service. En raison de ce coût, qui peut varier selon le type et la complexité de l'image, il faut se demander quels avantages pourraient découler de l'utilisation de divers genres d'images sur Télidon. Nous devons posséder une idée plus précise de la façon dont fonctionnent les images visuelles et de l'apport des images au processus de communication en général. Le présent rapport cherche à répondre à deux questions fondamentales: 1) Qu'est-ce que les recherches actuelles dans les domaines de la perception et de la cognition peuvent nous apprendre sur notre façon de comprendre et d'utiliser les images visuelles? 2) Quelles sont les répercussions de ces théories sur la façon de traiter et d'utiliser des images et des graphiques sur Télidon? En d'autres mots, qu'est-ce que les théories de la compréhension des images peuvent nous enseigner sur l'exploitation des avantages de Télidon en tant que moyen graphique?

Obtenir réponse à des questions aussi imprécises n'est pas une mince affaire. Pour pouvoir tirer parti des réponses, il faut ramener les nombreuses questions que pose la communication visuelle à un

ensemble de thèmes qui, tout en étant plus restreints, n'en demeurent pas moins pertinents. Notre examen de la documentation existante a fait ressortir trois thèmes. Malgré leur portée assez vaste, ces trois thèmes fournissent une structure d'organisation du présent rapport et chacun d'eux fait l'objet d'un chapitre.

1) Une image sur Télidon vaudra-t-elle mille mots?

Le premier chapitre examine d'emblée un argument, souvent invoqué par rapport à Télidon, selon lequel les images seraient supérieures aux mots parce qu'elles sont un moyen de communication plus naturel. L'étude de cette question débouche sur une question plus importante. Y a-t-il des différences fondamentales entre la façon dont les mots et les images transmettent de l'information? Pour répondre à cette question, nous avons tenté de déterminer s'il est possible de remplacer un énoncé verbal par une image. Le chapitre présente aussi quelques notions théoriques élémentaires qui soulignent la souplesse du processus d'interprétation visuelle. L'exposé de ces théories devrait nous aider à mieux comprendre pourquoi les images posent des problèmes dans certains secteurs de la recherche Télidon, notamment la description et le classement automatiques des images aux fins de recouvrement par ordinateur et l'élaboration d'index iconiques. Le chapitre se termine par un examen des recherches dont se dégage la véritable question concernant Télidon; il ne s'agit pas de savoir comment substituer des images aux mots, mais plutôt de déterminer comment la signification des mots peut dépendre d'un contexte pictural, comment la signification d'une image peut faciliter la compréhension d'un texte et enfin, comment la signification visuelle peut interagir avec la signification verbale pour engendrer de nouvelles connaissances.

2) Lire les images: le réalisme est-il ce qu'il y a de mieux?

Le chapitre II commence par demander s'il faut apprendre à lire les images un peu comme on apprend à lire l'écrit: quiconque peut voir le monde visuel peut-il voir le monde représenté par une image? N'apprend-on pas plutôt à lire les images comme on apprend un code symbolique, le processus dépendant alors du contexte, de la culture et des conventions? Vient ensuite une exploration des répercussions sur Télidon des recherches théoriques et empiriques qui indiquent que les images s'écartant du réalisme photographique, par exemple les dessins au trait, les bandes dessinées et les caricatures, sont plus facilement et plus rapidement reconnaissables que des photographies réalistes. Nous examinons l'idée voulant que l'efficacité des bandes dessinées et des caricatures soit due au fait que ces dessins reflètent la nature schématique de nos concepts intérieurs. Nous avançons plusieurs hypothèses pour montrer comment cette constatation pourrait permettre d'exploiter les avantages de Télidon en tant que médium pictural: c'est-à-dire sa capacité de produire des graphiques schématiques de grande qualité. Nous examinons en particulier les effets pondérateurs possibles, en termes de satisfaction de l'utilisateur et de facilité de compréhension, de variables comme les délais de transmission, le degré de réalisme pictural, l'objectivité que les gens attribuent aux diverses sortes d'images et la configuration spatiale de l'information.

3) L'image au service de la pensée: des schèmes graphiques pour véhiculer des idées complexes

Dans ce dernier chapitre, nous explorons quelques questions liées à l'utilisation d'images graphiques simples pour véhiculer des idées complexes. Nous examinons d'abord quelques-uns des processus psychologiques qui interviennent dans l'assimilation et l'utilisation des cartes, un des types d'images susceptibles de jouer un rôle

important pour Télidon. Nous soulignons les différences qu'il peut y avoir d'une personne à l'autre dans l'aptitude à se servir d'une carte. Nous constatons que les cartes très schématisées qui s'écartent de la réalité peuvent être plus faciles à comprendre parce qu'elles donnent les rapports fonctionnels entre les entités géographiques. Nous jetons aussi un coup d'oeil sur une expérience récente qui visait à remplacer la carte ordinaire sur papier par une "carte cinématographique" dynamique rendue possible par l'union des ordinateurs et des vidéodisques optiques. Nous nous interrogeons ensuite sur les raisons de la popularité des schémas graphiques simples (diagrammes de Venn, cases, flèches, etc.) dans des contextes scientifiques et pédagogiques. Leur popularité tiendrait à leur capacité de nous aider à saisir les propriétés génériques et les rapports sous-jacents qui ne peuvent être observés directement dans les caractéristiques superficielles des phénomènes physiques. Malgré l'utilité des schémas graphiques statiques, nous croyons que les graphiques dynamiques animés par des moyens informatiques pour montrer des transformations sont encore plus utiles à la pensée constructive et à la résolution de problèmes. Nous faisons, enfin, une distinction entre mouvement et séquence, convaincus que le fait d'être sensibles à cette distinction pourrait aider les concepteurs de systèmes à concevoir des techniques permettant de présenter des séquences visuelles dynamiques et de réaliser quand même des économies de largeur de bande d'émission.

Voilà donc, en résumé, les principaux thèmes qui sont développés dans notre rapport. Le lecteur pourra trouver abusive mon application de certaines théories cognitives à Télidon. Je ne m'en excuse pas. Au stade actuel de la recherche, nos connaissances concernant l'impact social et psychologique de nouveaux systèmes de communication sont encore si minces que les chercheurs en comportement ont véritablement besoin de quelques théories unificatrices pouvant orienter leurs efforts ou, tout au moins, les pousser à tirer certaines questions au clair,

même s'il faut pour cela prendre quelques libertés. Ce rapport soulève aussi de nombreuses questions qui restent sans réponse et il est muet sur beaucoup d'autres, fort intéressantes, qui concernent les images et Télidon (par exemple, le rôle de la couleur). On ne peut voir à tout dans un seul rapport!

En bref, le rapport vise non pas seulement à présenter une simple revue de la documentation existante mais à proposer un point de vue: en avançant certaines hypothèses, nous souhaitons favoriser (et peut-être provoquer) l'approfondissement de notre compréhension du rôle des représentations graphiques sur Télidon.

CHAPITRE I

UNE IMAGE SUR TÉLIDON VAUDRA-T-ELLE MILLE MOTS?

Imaginez le fournisseur d'information (FI) qui doit constituer une base de données Télidon sur un sujet quelconque. Au départ, il est aux prises avec deux problèmes. Il doit d'abord concevoir la structure représentationnelle qui saisira les rapports conceptuels entre les éléments contenus dans la base de données, et ensuite choisir le meilleur moyen de traduire ces concepts et rapports logiques pour qu'ils soient rapidement compris. Le message serait-t-il mieux exprimé sous forme de textes, d'images de différentes sortes ou encore par une combinaison des deux? Comme nous l'avons vu au chapitre précédent, le choix d'un moyen de communication -- par exemple, l'utilisation d'une certaine image, d'un graphique ou d'un énoncé verbal -- peut aussi être une question de coût, en termes de temps, de travail et même de mémoire informatique. Il est donc important de tenir compte des effets pondérateurs que pourraient comporter les différentes possibilités de codage de l'information offertes au FI en regard des avantages qu'en retirera l'utilisateur (soit la facilité ou la rapidité de compréhension, la clarté, la force attractive ou autre). Pour mieux comprendre la nature de ces facteurs de pondération, il convient tout d'abord d'élucider certaines questions d'ordre général concernant le mécanisme de compréhension des idées véhiculées par les mots et par l'image.

1.0. Un argument courant concernant les images et les mots

Examinons d'abord un argument que l'on entend souvent lors de discussions sur le rôle de l'image et du texte dans les systèmes vidéotex. Le raisonnement se résume comme suit. Nous vivons à l'époque du visuel. L'image est omniprésente: la publicité, le cinéma, les revues, les affiches, etc. Les gens, et surtout les enfants, n'ont aucune difficulté à comprendre les images. Les images sont en quelque sorte des véhicules "naturels" de communication car elles nous

renvoient directement à leurs référents; en d'autres termes, elles exploitent notre connaissance des "invariants" perceptuels qui caractérisent les objets et les événements dans notre environnement naturel (Gibson, 1966, 1971; Kennedy, 1974). La langue écrite et parlée, par contre, reposerait sur des codes symboliques conventionnels, propres à une culture donnée; ce sont des signes n'ayant que des rapports arbitraires avec les choses qu'ils représentent et, dans le cas de l'écriture, ils ne peuvent être assimilés qu'après un long processus d'apprentissage. On considère généralement que le langage interpose une plus grande distance entre le récepteur et le référent, constituant, de ce fait, un mode de communication moins direct et moins naturel que les images. Même si j'en ai quelque peu simplifié les termes ici, je ne crois pas exagéré de soupçonner un tel raisonnement d'être à l'origine de certaines propositions concrètes, faites au cours des dernières années, où l'on invoque la supériorité des images sur les mots dans différents contextes pédagogiques, scientifiques et sociaux. Citons, à titre d'exemple: le désir de créer un langage iconique "pur", exempt de symboles linguistiques, qui serait compris de tous, indépendamment de leur langue et de leur appartenance culturelle (pour une analyse des différents aspects de cette question, voir Modley, 1947; Mead and Modley, 1968; Kolers, 1969; Bliss, 1965); un plaidoyer en faveur d'une revitalisation du processus d'éducation par l'utilisation d'images et de graphiques -- notamment, à l'aide d'ordinateurs -- au lieu de descriptions verbales (Huggins and Entwisle, 1974, Ch. V); la nécessité de concevoir des interfaces homme-machine qui favorisent une relation directe et intime dans l'interaction homme-machine par l'utilisation de symboles non linguistiques (Johnson, 1970; Huggins, 1971).

Devant un tel engouement pour le potentiel communicatif de l'image par opposition aux symboles linguistiques, il ne serait pas étonnant qu'un fournisseur d'information d'un système Télidon se demande "Est-ce-que ce que j'ai à dire serait mieux exprimé en images qu'en mots?" Bien

sûr, comme nous le verrons, il y a peu de chances qu'on trouve une réponse toute faite à la question de savoir ce qui est "mieux". Le choix d'utiliser un type particulier d'images ou de graphiques plutôt qu'une description linguistique sera probablement dicté par des circonstances spécifiques et non par une recette. Cela signifie que l'on tiendra compte de la nature du "message" que l'on veut communiquer, de la clientèle visée et d'une foule d'autres facteurs.

À la question de savoir si les images sont plus efficaces que les mots (ou inversement) et dans quelles situations cela se vérifie, la réponse sera relativement complexe. On pourrait départager plus clairement les mérites relatifs des images et des mots si, au départ, au lieu de chercher à déterminer ce qui est "mieux", on se posait une question plus fondamentale, soit "Ce que j'ai à dire est-il exprimable en images?" En d'autres termes, il est peut-être opportun de se demander s'il est vrai qu'une image ou une séquence d'images peut remplacer un énoncé linguistique. Chercher la réponse à cette question d'équivalence pourrait non seulement nous révéler quelques différences qualitatives importantes entre la fonction communicative des images et celle des mots, mais aussi nous mettre indirectement sur la voie de ce que je crois être la véritable question pour Télidon, c'est-à-dire comment l'intervention réciproque des mots et des images introduit des contraintes dans le processus de compréhension.

1.1. Une image peut-elle remplacer un énoncé verbal?

Voyons, à l'aide d'un simple exemple donné par E.H. Gombrich (1972), s'il est juste d'affirmer que les énoncés verbaux sont traduisibles par leurs équivalents picturaux. Prenons la phrase -- "Le chat est sur la carpette" -- qui décrit un simple état de choses. Il est relativement facile de s'imaginer la situation décrite par cette phrase. En effet, on passe rapidement de la formation d'une "image mentale" d'un chat sur une carpette à la création d'une image réelle qui, dans le cas

qui nous occupe, est un croquis représentant l'énoncé "Le chat est sur la carpette." (Voir figure 1.0). Mais l'énoncé et le dessin ont-ils vraiment la même signification? À première vue, on serait porté à le croire. Le croquis permet évidemment à un observateur, tout au moins de culture occidentale, d'identifier les objets familiers "chat", "carpette" et la relation spatiale "sur" que décrit l'image. Toutefois, si on regarde plus attentivement la phrase en question, on voit certaines différences intéressantes. Par exemple, la phrase montre clairement qu'il s'agit de "le chat" (un individu) et non de "un chat" (un membre de l'espèce). Comment l'image pourrait-elle, à elle seule, exprimer cette distinction? L'article défini "le" introduit des contraintes importantes dans le processus d'interprétation; il commande au système conceptuel d'appliquer et d'associer le nom "chat" à un individu en particulier à l'intérieur d'une situation que l'on connaît déjà ou qui sera bientôt identifiée. Nous savons maintenant que des mots simples comme "le" et "un" jouent un rôle important dans le processus de compréhension, grâce à des recherches menées récemment sur l'intelligence artificielle et dont l'objectif est d'amener un ordinateur à comprendre le langage naturel. Par exemple, le programme de compréhension du langage SHRDLU, présenté par Winograd (1972), détermine le sens de l'article "le" à l'aide de procédures ou programmes qui utilisent non seulement les règles syntaxiques de l'anglais mais aussi sa capacité de mémorisation d'événements récents et de connaissance du monde -- dans le cas du programme SHRDLU, la connaissance des changements dans les relations spatiales d'un "monde de blocs" -- pour déterminer si un nom se rapporte à un individu en particulier qui lui est déjà familier (noeud secondaire dans la mémoire) ou à son générique (noeud primaire). Notre dessin, en lui-même, ne comporte aucun élément qui indique au système conceptuel s'il s'agit "du" ou "d'un" chat. Cela vaudrait également pour une photographie d'un vrai chat sur une carpette car, même là, nous ne saurions pas s'il désigne un chat en particulier ou les chats en général.

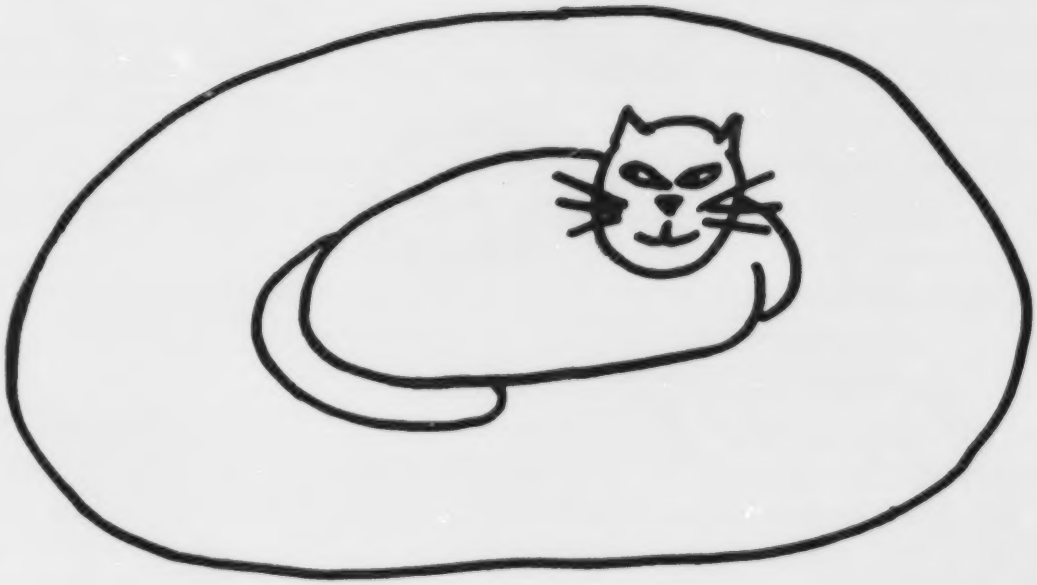


Figure 1.0. "Le chat est sur la carpette"

Si déjà le dessin de la figure 1.0 arrive mal à décrire les fonctions qu'expriment "le" et "un", il serait nettement déclassé, face aux pouvoirs descriptifs du langage, si l'on voulait rendre par des équivalents picturaux des phrases telles que "le chat se lèvera de la carpette", "il n'y a pas de chien avec le chat sur la carpette", ou encore "si un chien entre dans la pièce, le chat quittera la carpette". D'après Gombrich, des phrases de ce genre révèlent le véritable miracle du langage, qui nous permet de décrire des situations présentes, passées, futures et même conditionnelles. De plus, la langue contient quantité de mots "logiques" comme "si", "quoique", "mais", "lorsque" et autres, qui semblent relever intégralement de la capacité que nous avons de faire des déductions et des inférences à propos des situations. Quelle est l'image qui pourrait rendre la préposition "si" (mais voir Arnheim, 1969, ch. 10)? Bien sûr, on peut toujours prétendre qu'à défaut d'une seule image qui puisse accomplir de telles prouesses, il pourrait être possible de concevoir une séquence d'images capable de décrire des états passés, futurs et conditionnels. Aux défenseurs de cette thèse, nous pourrions proposer un exercice salutaire qui consisterait à rendre purement en images, sans recours à aucun symbole linguistique, la phrase "le chat va se lever de la carpette". Parions qu'un tel effort aura vite fait de soulever leur admiration pour l'efficacité descriptive du langage verbal.

1.1.1. Une image vaut mille mots, mais lesquels?

Il importe de faire ici une mise au point sur la question que nous soulevons à propos des images et du langage. Nous ne prétendons pas qu'une image est nécessairement inférieure à un énoncé verbal, pas plus que nous n'affirmons que les images sont incapables de communiquer des renseignements utiles sur les objets, les situations et les événements. Ce que nous voulons faire ressortir, c'est qu'il semble y avoir effectivement un problème fondamental lorsqu'on utilise une image au

lieu d'un énoncé verbal. Le problème n'est pas qu'une image ne vaut pas mille mots. Une image peut effectivement valoir mille mots mais le vrai problème est que, en l'absence d'un certain contexte limitatif, souvent nous ne savons pas lesquels. En d'autres termes, toute image, même si elle est aussi simple que la figure 1.0, peut donner lieu à un nombre infini d'interprétations. Les images, à elles seules, ne comportent aucun moyen simple de guider le récepteur vers le sous-ensemble de propositions visé par l'émetteur.

En résumé, nous avons soulevé dans la présente section la question de savoir si une image pourra jamais remplacer un énoncé verbal. Nous avons constaté que l'un des problèmes reliés aux images en tant que média "autonome" réside dans le fait qu'elles ne possèdent aucun moyen simple de limiter le nombre infini d'interprétations auxquelles elles se prêtent. Dans la section suivante, nous examinerons plus à fond ce problème de polysémie des images en exposant certains points de vue sur le mode de fonctionnement de la perception en général. Ces points de vue, qui soulignent la souplesse du processus perceptuel et sa dépendance du contexte, nous aideront à mieux comprendre pourquoi les images posent de sérieux problèmes dans un certain nombre d'applications liées à Télidon. Par exemple, comment classer des images ou des scènes en vue de la recherche documentaire par ordinateur, ou encore qu'implique la mise au point "d'index iconiques"? Ces exemples montreront à quel point le contexte est important dans le processus de compréhension et attireront notre attention sur certaines questions importantes concernant l'utilisation de l'image et du langage sur Télidon. Nous verrons comment le langage donne le contexte nécessaire à la compréhension des scènes, comment une image restreint le champ d'interprétation du texte et, enfin, comment les images et les mots interagissent pendant le processus de compréhension.

1.2. Polysémie des images: Les "schémas mentaux" dans la perception

Nous avons cherché à démontrer que le sens véhiculé par une image -- même une image très simple comme la figure 1.0 — peut être plus ou moins clair. Toute image, en tant que moyen de communication, présente un problème puisque, seule, elle ne peut diriger le récepteur vers le sens particulier que, parmi le nombre infini de sens possibles, l'émetteur a voulu lui donner. On peut se demander toutefois si ce problème de polysémie est particulier aux images picturales ou s'il constitue plutôt une caractéristique inhérente à l'interprétation visuelle en général. Ce problème de polysémie se présente-t-il, par exemple, dans la perception quotidienne des événements et des objets du monde visuel? Il est certain qu'en nous déplaçant dans le monde visuel nous éprouvons rarement des difficultés à identifier les objets, les gens et les événements de notre entourage. Cela demeure vrai même si, comme le soulignent certains théoriciens classiques de la perception (voir annexe A), une forme bidimensionnelle donnée qui est projetée sur la rétine peut avoir été produite par n'importe lequel parmi un nombre infini d'objets de formes différentes observés sous divers angles et à différentes distances de l'observateur. De plus, l'information perçue par l'appareil sensoriel se modifie constamment au fur et à mesure que nous nous déplaçons dans le monde. Cela est dû aux changements de point de vue, à l'interposition d'autres objets entre l'observateur et l'objet observé, aux changements momentanés d'éclairage et à d'autres facteurs. Un des grand défis pour la théorie est d'expliquer comment nous pouvons maintenir une perception stable du monde en dépit de ces déplacements et de ces changements incessants.

Il y a une méthode d'aborder ces problèmes de polysémie et de flux constant qui consiste à en nier l'existence. C'est la position adoptée par J.J. Gibson et ses disciples (voir analyses récentes de la position gibsonienne, dans Shaw et Bransford, 1977; Kennedy, 1975), qui soutiennent que nos systèmes perceptuels rencontrent rarement des

problèmes d'ambiguïté de perception dans la vie de tous les jours du fait que le système visuel est "réglé" pour "capter" dans le champ optique des invariants informationnels importants de "second ordre" qui permettent l'identification automatique et sans ambiguïté des objets, des gens et des événements dans l'environnement naturel. Au lieu de considérer les mouvements de tête et les changements de position du corps comme des problèmes, les gibsoniens les interprètent plutôt comme des sources d'information utile "temporellement variable" pour le système visuel -- les structures optiques potentiellement descriptibles par les lois de la "physique écologique" (gradients de texture et éclairage, angles et lignes d'occlusion, etc.) qui éliminent les sources d'ambiguïté apparente inhérentes aux images fixes. Une conception gibsonienne de la perception, parfois appelée "réalisme direct" minimise donc les problèmes que soulèvent l'ambiguïté, l'expérience du vécu et le rôle des inférences mentales dans la perception — éléments-clés de ce qu'on appelle parfois l'approche "cognitiviste" ou "constructiviste". Une analyse détaillée du débat opposant constructivisme et réalisme direct dépasse le propos du présent exposé, mais l'annexe A donne un bref aperçu de la question. Il suffira de mentionner ici qu'en raison de son refus de reconnaître l'apport de la mémoire, de l'apprentissage et des inférences mentales à la perception, nous estimons l'approche gibsonienne un peu moins utile que la théorie constructiviste en tant que schème de réflexion pour le problème des images Télidon (voir néanmoins la contribution de Kennedy, 1974, au ch. 2 de la présente étude).

1.2.1. Schémas et transformations

Sans entrer dans les détails, abordons brièvement une certaine théorie constructiviste de la compréhension perceptuelle qui peut nous éclairer sur le problème de la polysémie des images. Cette théorie part du principe que la perception et la compréhension en général sont des

processus de "résolution de problèmes" ou de "théorisation" par lequel le système conceptuel cherche à expliquer les données brutes qui atteignent les sens en élaborant une "description" -- une sorte de théorie ou de modèle mental du champ sensoriel qui interprète ces données en les rattachant à un ensemble de concepts sous-jacents (Bregman, 1977, 1979; Rock, 1975; Minksy, 1975). Selon Bregman (1977), les descriptions (théories) des données sensorielles pourraient être élaborées à partir de deux principaux types de concepts ou unités mentales: les "schémas", que Bregman appelle "idéaux", et les "transformations". Un schéma ou idéal peut être défini comme étant une structure mentale qui encode la connaissance que le cerveau possède de toute régularité ou modèle cohérent du vécu; ces structures s'apparenteraient à des "formules" ou "stéréotypes" qui sont le rendu de notre connaissance prototypique des objets, des événements et des situations. Tandis que les schémas ou idéaux traduisent la connaissance prototypique des entités, les transformations sont des concepts qui rendent compte des changements ou modifications qui peuvent se produire dans la forme prototypique des schémas. Les concepts transformationnels jouent un rôle primordial car il y aura toujours un certain décalage entre l'idéal mental ou le prototype qui représente une entité dans le cerveau et la façon exacte dont cette entité peut se produire dans une scène ou situation spécifique de la réalité. Dans le cas de la perception visuelle, ces écarts peuvent être causés par l'angle visuel particulier de l'observateur, par la présence d'autres objets, par des changements d'éclairage ou par d'autres facteurs. On pourrait illustrer la façon dont opèrent les schémas et les transformations dans l'interprétation visuelle par un exemple fort simple, soit la capacité qu'a le système conceptuel d'expliquer une forme elliptique projetée sur la rétine en l'interprétant comme étant un schéma -- un cercle parfait -- qui a subi une transformation <u>oblique</u> due à l'angle de vision particulier de l'observateur. L'approche schématico-transformationnelle de la perception cherche à montrer comment le système conceptuel pourrait utiliser un ensemble fini de concepts généraux ou schèmes pour

comprendre une variété infinie de scènes et d'événements spécifiques. Cette idée n'est pas nouvelle en psychologie; on la retrouve dans les travaux de Bartlett (1932), Piaget (1947) et, plus récemment, dans la théorie des "frames" ou planches de Minksy. Elle constitue aussi une tentative d'alignement sur la théorie linguistique de Chomsky (1975) selon laquelle un ensemble fini de composantes d'une structure "profonde" combiné aux règles de la transformation permettrait de générer un nombre infini de phrases nouvelles au niveau superficiel.

Les schémas et les transformations interviennent également dans notre processus d'interprétation des événements temporels, comme on peut le voir dans la bande dessinée de la figure 1.1. On pourrait donner une description "objective" de ce dessin en ces termes: "En partant de la gauche, on compte quatre rectangles adjacents. Chaque rectangle renferme un point noir toujours plus gros que celui qui le précède. Le dernier rectangle est complètement noir." C'est là une description possible, quoique passablement profuse en tant que "théorie" des données sensorielles. Nos systèmes cognitifs opteront fort probablement pour une interprétation plus élégante et économique fondée sur des concepts transformationnels plus "profonds". Par exemple, au lieu d'y voir quatre images de quatre objets distincts et indépendants, la connaissance que nous avons des règles du changement perceptif, conjuguée à notre connaissance des règles de la lecture de bandes dessinées, engendre une représentation mentale du dessin où les formes séparées sont perçues comme un objet "idéal" <u>unique</u> qui subit un certain type de transformation. Ainsi, par exemple, on peut interpréter les changements de volume et de forme comme étant la représentation d'un seul objet qui se dirige vers un observateur immobile. Une telle transformation recèlerait des interprétations du dessin telles que "se trouver face à un taureau qui fonce" ou "recevoir une balle sur le nez". Le "mouvement" dans cet exemple est donc un concept tranformationnel qui, une fois inséré dans la représentation mentale du dessin, permet au système conceptuel de percevoir un patron complexe -- les points isolés

de différentes grandeurs (y compris le rectangle noirci), comme des régularités sous-jacentes plus simples -- c'est-à-dire un objet idéal en mouvement.

L'approche schématico-transformationnelle de la compréhension visuelle peut nous aider à saisir pourquoi nous pouvons lire dans une image un nombre infini de propositions différentes. En effet, si la perception est en quelque sorte un mécanisme de résolution de problèmes par lequel le système conceptuel cherche à expliquer un patron visuel par l'élaboration d'une théorie ou d'une description fondée sur un ensemble fini de schémas sous-jacents, le même patron ou "agencement de données" peut aussi être expliqué par des théories concurrentes, c'est-à-dire par des combinaisons différentes de concepts sous-jacents. Le choix de la combinaison de concepts qui sera appliquée à une situation donnée sera dicté par les objectifs et les attentes du récepteur, le contexte et d'autres facteurs. La capacité que nous avons d'appliquer plus d'une description à un même patron ou événement -- surtout comme conséquence de changements intervenus dans le contexte -- confère de la souplesse à nos systèmes cognitifs et se trouve vraisemblablement au coeur même de nos facultés créatrices et de notre aptitude à résoudre des problèmes.

C'est lorsque les contraintes contextuelles sont minimes et que les éléments visuels d'une image sont relativement abstraits qu'on peut vraiment constater la souplesse du processus d'interprétation. Examinons, par exemple, certaines des interprétations du dessin de la figure 1.1 fournies par des collégiens à qui on a demandé à quel événement familier ce dessin leur faisait penser (voir Mills, 1980a et 1980b pour une analyse approfondie des processus psychologiques qui interviennent dans la formulation de descriptions verbales de bandes dessinées "abstraites").

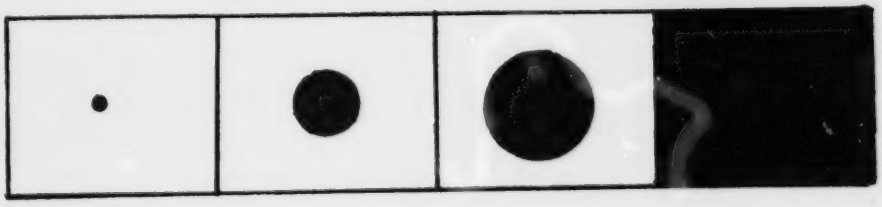


Figure 1.1. Une bande dessinée

Extrait de: Mills (1980)

1) on me lance une pierre;
2) je reçois une balle sur le nez;
3) je regarde une tête de clou de différentes hauteurs;
4) je vois une planète grossir au fur et à mesure que je m'en approche dans un vaisseau spatial;
5) c'est un rond dans l'eau provoqué par un caillou jeté dans un étang;
6) c'est un problème insoluble auquel on doit se résigner.

Même s'il nous est impossible de faire ici une analyse détaillée de ces interprétations, le lecteur comprendra facilement comment différentes combinaisons de schémas et de transformations (y compris les hypothèses des observateurs sur leur propre rôle dans l'événement) arrivent à expliquer les formes et les patrons de changement dans le dessin. Par exemple, le patron de changements de la figure 1.1 peut être interprété non seulement comme un seul objet se dirigeant vers un observateur stationnaire (interprétation 1) mais aussi comme un observateur en mouvement qui s'approche d'un objet immobile -- soit une autre sorte de transformation concrétisée dans l'interprétation "je vois une planète grossir au fur et à mesure que je m'en approche dans un vaisseau spatial". D'autres interprétations sont possibles à partir de combinaisons, de schémas et de transformations différentes.

Il est vrai que, dans cet exemple, la simplicité et le caractère abstrait du dessin ainsi que l'absence de limites contextuelles ouvrent la voie à un vaste éventail de descriptions possibles -- et qui peut dire laquelle est la meilleure? Nous croyons, néanmoins, qu'il n'y a pas de différence fondamentale entre l'interprétation d'un dessin "abstrait" comme celui de la figure 1.1 et l'interprétation d'images ordinaires dont les éléments représentationnels cherchent à décrire l'aspect visible d'objets reconnaissables. Dans les deux cas, le système conceptuel doit faire appel à des concepts ou règles (schémas et

transformations) qui permettent d'établir une représentation corrélative entre le domaine représentationnel (lignes, taches de couleur sur papier) et le domaine cible (événements reconnaissables, objets en trois dimensions, etc.). Comme nous le verrons plus loin, pour chaque élément visuel -- image réaliste ou dessin abstrait -- il y a différentes représentations possibles des mêmes données selon les changements de contexte, y compris les objectifs et les attentes du récepteur. Ce fait a été dépeint de façon amusante par le dessinateur Saul Steinberg qui, en une seule image, montre comment un élément graphique simple -- la ligne -- peut prendre des significations très différentes si on en modifie le contexte (voir figure 1.2).

1.3. Descriptions d'images destinées aux machines

Les implications d'une approche "descriptionnelle" à la perception visuelle qui met en évidence la nature non rigide et la dépendance du contexte du processus d'interprétation ne sont pas purement théoriques. Dans le domaine de la recherche appliquée, les travaux visant à mettre au point des systèmes de recherche documentaire ayant des éléments de nature picturale constituent un exemple des problèmes de souplesse et de contexte inhérents à l'interprétation visuelle auxquels il faudra trouver une solution d'ordre pratique. (Firschein et Fischler, 1972, 1969; Fischler, 1969; Narasimhan, 1969, Clowes, 1969). Ces travaux permettraient à des utilisateurs de questionner un système sur l'existence d'une image en particulier ou d'un ensemble d'images et même sur un sous-ensemble de propriétés que l'image peut contenir (qu'on peut présumer y être contenu). Par exemple, un tel système pourrait avoir à répondre à une question comme celle-ci, fondée sur la description d'une image qui y est stockée: "Est-ce qu'on y trouve une voie navigable?"

Figure 1.2 Dessin de Steinberg.

Source: New Yorker Magazine (1954).

Firschein and Fischler (1972) ont comparé le problème de mise au point d'un système de recherche documentaire qui utilise des images comme bases de données au problème que pose la préparation d'une "encyclopédie" -- c'est-à-dire comment structurer un vaste ensemble de connaissances de sorte que la consultation en soit facile pour différents types d'utilisateurs ayant différentes fins. La création d'une telle encyclopédie fondée sur l'information issue d'une image n'exige rien de moins que de trouver une façon de représenter le "message" d'une image donnée au moyen d'une structure symbolique quelconque -- par exemple, un réseau sémantique qui utilise un ensemble fini de noeuds et de liaisons pour rendre, sous une forme canonique, les relations conceptuelles entre les éléments de l'image (par exemple X est à l'ouest de Y (emplacement), X est un Y (inclusion de classe), X se prolonge à travers Y (opérations) etc.).

Pour produire des représentations encyclopédiques d'images, on peut, entre autres méthodes, demander à un groupe de personnes ayant des formations différentes, des objectifs et des degrés de compétence ou d'autres caractéristiques dissemblables de décrire verbalement une image donnée (par exemple, une photo aérienne d'un certain territoire). On fait alors un "recodage" de ces descriptions, selon un ensemble de règles, sur le réseau sémantique sous-jacent. Les structures sémantiques issues des descriptions données par différents sujets sont ensuite combinées pour former une seule représentation "composite" qui, en théorie, pourrait servir à répondre à des questions portant sur l'image selon différents points de vue ou contextes (par exemple, géographie, géologie, politique ou renseignement militaire). Outre le problème que pose la multiplicité éventuelle des contextes ou des points de vue (différents schémas sous-jacents), la création d'une telle encyclopédie "composite" à partir de descriptions humaines se heurte à d'énormes difficultés. Par exemple, une personne pourra décrire une partie d'une image comme étant "une maison avec une piscine et un

terrain de tennis" -- ce qui constitue un commentaire relativement objectif d'un groupe d'objets contenu dans l'image. Pour une autre personne, toutefois, ces mêmes objets peuvent représenter "une maison luxueuse" (Firschein and Fischler, 1972). Ceci constitue manifestement une démarche déductive qui transpose l'information explicite de la photo. Le fait qu'une image comportant un ensemble fini d'objets puisse engendrer un nombre pratiquement illimité de nouvelles inférences (nouvelles descriptions), semble poser un obstacle majeur à la création d'une représentation "encyclopédique" de la signification d'une image. Une analyse des problèmes (et des solutions) que posent la description et l'extraction automatiques des scènes dépasse la portée de la présente étude et la compétence de l'auteur. Ces questions seront toutefois d'une importance évidente pour l'avenir de Télidon. On peut facilement s'imaginer des situations où des gens voudront utiliser Télidon pour consulter diverses sortes d'images ainsi que pour obtenir des renseignements sur des images sans nécessairement regarder l'image elle-même.

1.4. Les images peuvent-elles servir d'index?

Jusqu'à maintenant, nous nous sommes demandé si une image, à elle seule, pourrait remplacer un énoncé verbal. Nous avons cherché à démontrer que l'utilisation d'une image comme substitut pose un problème fondamental qui trouve ses racines dans la nature même de l'appareil conceptuel humain. Ce dernier est un instrument extrêmement souple qui peut s'inspirer d'un ensemble fini de concepts (schèmes) et les agencer selon différentes combinaisons pour engendrer un nombre infini de "descriptions" (représentations mentales) différentes d'un patron visuel donné, dont la source est soit une surface picturale, soit un objet quelque part dans le monde. On trouve toutefois des exemples de communication picturale où le système conceptuel ne puise pas indéfiniment de nouveaux sens, mais où la signification des images est

claire, voire où le langage verbal lui-même est peu utile. Le meilleur exemple de ce type de communication peut être observé dans les événements internationaux comme les Jeux olympiques où l'on utilise des pictogrammes comme index de sports pour indiquer à l'observateur, indépendamment de sa langue maternelle et par moyen visuel seulement, de quel événement sportif en particulier il est question. Ces index picturaux témoignent souvent d'une grande ingéniosité, notamment par l'utilisation de formes abstraites qu'on reconnaît facilement, malgré leur simplicité, comme la représentation d'un athlète, d'un objet et d'une action. Nous examinerons un peu plus loin certaines de leurs propriétés.

E.H. Gombrich nous rappelle que l'utilité des symboles iconiques tels les panneaux de signalisation olympiques dépend en grande partie de la clarté du contexte dans lequel ils sont utilisés. Lorsque le contexte est confus ou obscur, l'utilisation d'images comme index ou étiquettes peut avoir des conséquences qui, même si elles sont parfois comiques, sont fort éloignées du but recherché. Gombrich raconte:

> Il y a quelques années, on rapportait dans les journaux que des émeutes avaient éclaté dans un pays en voie de développement à cause d'une rumeur à l'effet qu'on vendait de la chair humaine dans un magasin. On découvrit que cette rumeur était reliée à des boîtes de conserve dont l'étiquette montrait un garçon au large sourire. Ici, c'était le changement de contexte qui avait engendré la méprise: en règle générale, une image de fruits, de légumes ou de viande sur une boîte de conserve indique effectivement son contenu; nous ne faisons jamais la même déduction quand c'est l'image d'un être humain qui est sur la boîte parce qu'au départ, c'est une possibilité que nous n'envisageons même pas. (Gombrich, 1972, p. 87).

Les signes olympiques ne prêtent pas à une telle confusion car le contexte est clair et les panneaux peuvent, par conséquent, très bien servir d'index même sans l'aide de mots. Cependant, même lorsque le contexte est clair, on peut se demander si certains types d'images fonctionnent mieux que d'autres en tant qu'index. Cette question est pertinente compte tenu de l'utilisation possible d'une certaine forme d'indexation picturale sur Télidon. Faisons une brève digression pour étudier une question difficile que nous n'avons pas encore abordée. Quelles fonctions remplissent les images en général et est-ce que certaines fonctions sont mieux servies par certaines propriétés visuelles? Cette question a déjà été traitée par Rudolph Arnheim (1969), dont nous nous sommes largement inspiré.

1.4.1. Icônes, signes et symboles

Arnheim fait une distinction utile entre les propriétés physiques d'une image et la fonction communicative qu'elle peut remplir. Cette distinction entre forme et fonction peut être mieux perçue si l'on montre comment la même image peut remplir différentes fonctions, parfois simultanément. D'après Arnheim, une image peut jouer le rôle d'icône, de symbole ou de signe (employé ici dans son sens restreint de "signal"). Prenons l'exemple d'un simple triangle. Un triangle a valeur d'icône lorsqu'il fait allusion à une montagne, de signe lorsqu'on l'utilise pour avertir d'un danger et de symbole lorsqu'il représente l'idée de hiérarchie. En résumé, une image joue le rôle d'icône, selon Arnheim, lorsqu'elle représente les choses à un niveau d'abstraction inférieur au sien. Ainsi, par exemple, on peut utiliser un triangle -- forme abstraite -- pour décrire une montagne. La "fonction iconique" est possible car triangle et montagne possèdent des caractéristiques structurales qui sont partiellement isomorphes (base, sommet). Mais un triangle est une forme schématique et simplifiée par

comparaison à une vraie montagne -- bien que les deux partagent certaines propriétés structurales. Une image sera un <u>symbole</u> si elle représente les choses à un niveau d'abstraction <u>supérieur</u> au sien. Par exemple, les caractéristiques structurales d'un triangle peuvent être utilisées non seulement pour décrire des formes physiques mais aussi pour représenter des idées, telle la notion logique de "hiérarchie". Selon Arnheim "les symboles confèrent une forme particulière à des catégories de choses ou à des constellations de forces". Une importante catégorie d'images dont nous parlerons plus loin, et qui remplissent des fonctions symboliques, sont les graphiques abstraits ou schématiques comme les formes d'ondes, les diagrammes de Venn, les flèches, les cases et autres symboles qui confèrent une forme visuelle à des patrons purement abstraits de forces, d'intensités, de tensions, de vecteurs, de chevauchements, etc. Enfin, une image joue le rôle de <u>signe</u> lorsqu'il existe entre elle et son référent un lien arbitraire, c'est-à-dire un lien simplement associatif, où les caractéristiques structurales de l'image ne comportent aucune correspondance isomorphe avec la structure des choses qu'elle suggère. On peut évidemment dire que les mots nous offrent l'exemple de signes <u>par excellence</u>. En effet, il n'y a aucune similitude physique entre le mot "chat" et l'animal qu'il représente. Il n'est pas toujours facile de déterminer si une image joue le rôle de symbole ou de signe. Par exemple, dire que notre triangle est un signe de danger, comme nous l'avons fait, est un peu délicat car un cercle pourrait tout aussi bien être signe de danger. Ce qui importe alors c'est que le cercle soit clairement distingué des autres formes et utilisé de façon constante. Mais il y a quelque chose dans le tracé d'un triangle qui le rend supérieur au cercle comme signe de danger, c'est-à-dire que les pointes et les contours abrupts du triangle semblent évoquer des qualités structurales abstraites à un niveau d'abstraction plus profond, de sorte qu'il sert simultanément de symbole et de signe.

Nous pouvons maintenant revenir à la question des index picturaux et voir si certains types d'images remplissent cette fonction d'indexation mieux que d'autres. Tout d'abord, quelles sont les caractéristiques d'un indicateur efficace? Un bon indicateur, comme nous l'avons déjà laissé entendre, est celui qui sélectionne rapidement et sans ambiguïté son référent parmi un vaste éventail de possibilités. À prime abord, on serait porté à dire que cette fonction sélective serait mieux servie par des images très réalistes. Mais est-ce vrai? Les recherches menées sur la description automatique de scènes, que nous avons abordées dans la section précédente, ne montraient-elles pas que même une photo réaliste peut donner lieu à un nombre infini de descriptions possibles? La difficulté peut tenir, en partie, du fait que plus une image est réaliste, plus elle contient d'éléments susceptibles d'accaparer l'attention de l'observateur. Arnheim décrit comment ce problème s'est présenté lors d'une expérience qui consistait à utiliser des images réalistes en guise de signaux routiers. Imaginez qu'on utilise l'image concrète d'un escargot pour indiquer à l'automobiliste de ralentir. Même si l'on considère généralement l'escargot comme un animal qui se déplace lentement, l'image réaliste d'un escargot peut suggérer d'autres qualificatifs tels que limoneux, petit, fragile, patient. Le contexte routier peut aider à concentrer l'attention de l'observateur sur le choix à faire, mais l'image elle-même n'effectue pas cette opération de convergence.

L'utilisation d'images réalistes comme index pose un autre problème qui est relié à la structuration de nos systèmes internes de classification. Un certain objet peut appartenir à un grand nombre de superensembles de sorte que l'image d'un objet, même s'il est facilement reconnaissable, ne comporte aucun moyen de spécifier son superensemble précis. L'image d'une clé anglaise accompagnant un ensemble de rubriques affichées sur Télidon peut indiquer à l'utilisateur qu'il existe un bloc d'information sur la mécanique en général. Mais, comment l'utilisateur peut-il savoir si cette clé veut dire "plombiers",

"quincailliers" ou "mécanique automobile"? L'image de la clé attire l'attention, mais elle peut mettre sur une fausse piste l'utilisateur qui ignore lequel des superensembles est ainsi désigné. Le fournisseur d'information Télidon se trouve donc devant un sérieux dilemme car il doit penser non seulement aux éventuels mérites d'une image concrète et accrocheuse, mais aussi au fait qu'une telle image risque de faire perdre du temps à l'utilisateur en l'orientant vers de mauvais superensembles. À moins que le contexte soit très clair, l'utilisation de mots pour désigner le superensemble serait peut-être le moyen le plus simple et le plus efficace d'indiquer une catégorie en particulier -- l'image qui l'accompagnerait serait alors redondante ou, pire encore, ralentirait le processus de recherche dans la base de données en attirant l'attention de l'utilisateur et en l'orientant sur une mauvaise voie. Bien sûr, ce sont là des questions empiriques qui pourraient être mises à l'épreuve.

On a déjà fait, sur la cognition, des travaux qui pourraient s'avérer utiles à notre réflexion sur les problèmes d'indexation -- linguistique ou picturale. Il s'agit de recherches menées sur la nature de la catégorisation humaine (Rosch, 1975; Rosch et Lloyd, 1978). Ces recherches ont montré, par exemple, que les membres d'une même classe ne sont pas tous également représentatifs de cette classe -- c'est-à-dire que certains sont plus prototypiques que d'autres. Ainsi, si l'on veut suggérer le superensemble "mobilier", il est préférable d'utiliser des termes comme "table" ou "chaise" plutôt que "sofa" ou "banc". Le même principe s'appliquerait lorsqu'on utilise des images de différents objets au lieu de mots pour suggérer des superensembles. On trouve un exemple pictural intéressant dans les recherches menées récemment sur la catégorisation dans le langage gestuel ASL (American Sign Language) -- langage purement "iconique" utilisé par les sourds (Newport et Bellugi, 1978; Bellugi et Klima, à l'édition). Ce langage fait appel à des gestes ou signaux manuels qui servent non pas à épeler des mots, mais

plutôt à tracer des "images dans l'air". Par exemple, le signe pour désigner une "guitare" en langage gestuel consiste à jouer d'une guitare imaginaire. Ces gestes manuels constituent un moyen de communication qui permet non seulement de décrire des objets concrets (chaise, guitare) mais aussi de rendre des idées abstraites (démocratie). Une des caractéristiques intéressantes du ASL est qu'il permet à un émetteur d'éveiller dans l'esprit du récepteur, à l'aide de gestes purement iconiques, la notion d'un "superordonné", comme par exemple "mobilier" ou "instruments de musique"? Comment parvient-on à représenter ces idées quand on sait qu'il n'y a dans ce langage gestuel aucun signe unique qui désigne un superensemble? La solution consiste à enchaîner plusieurs signes représentant séparément des objets qui sont des membres prototypiques de la classe. La seule différence est l'accélération du rythme de signalisation. Par exemple, pour dire "instruments de musique" en langage gestuel, l'émetteur fait d'abord le geste qui signifie "guitare" suivi rapidement par les signes conventionnels pour "piano" et "trompette", c'est-à-dire, les prototypes qui renvoient au superensemble. Il existe également des moyens d'indiquer visuellement le sous-ensemble (par exemple "piano droit") mais, pour des raisons d'espace, nous ne traiterons pas en détail de cette question.

Qu'est-ce que l'utilisation de signes dans ce langage gestuel nous enseigne, en termes pratiques, pour l'utilisation d'images comme index sur Télidon? Elle suggère tout d'abord qu'avant d'utiliser une image comme index, on doit s'assurer de son caractère "prototypique". En second lieu, est-ce que deux ou trois images d'objets prototypiques pourraient servir mieux qu'une seule à renvoyer à un superensemble? On pourrait peut-être faire d'autres découvertes précieuses pour l'utilisation d'images et de graphiques sur Télidon en étudiant la nature du langage gestuel. Il semble que les recherches menées sur la catégorisation humaine puissent être appliquées à la solution des problèmes d'indexation -- picturale ou verbale -- que pose Télidon.

Fig. 1.3. Pictogrammes: Jeux olympiques de Montréal.

Les index picturaux peuvent être utiles sur Télidon, mais est-ce que des images réelles feront toujours les meilleurs index? Pas nécessairement. Une image réaliste qui nous apparaîtra facilement comme étant la représentation d'un objet en particulier n'est peut-être pas le meilleur moyen d'indiquer rapidement les propriétés génériques qui caractérisent une catégorie d'information. Il se peut que, pour faciliter la perception du sens générique, il soit préférable d'utiliser des images abstraites et plus dépouillées. Pour indiquer une direction à suivre, nous dit Arnheim, une flèche stylisée est un meilleur symbole que ce "dessin réaliste de l'époque victorienne représentant une main avec les ongles, la manche et les boutons de manchette". Les images abstraites -- symboles -- ont le pouvoir de fixer notre attention sur les propriétés génériques avec plus de précision que ne le font les images réalistes. Toutefois, même si une image très abstraite met en relief des propriétés générales, elle a le désavantage de trop élargir le champ d'interprétation. Par exemple, notre triangle, bien qu'étant simple, abstrait et facilement identifiable, possède des caractéristiques structurales que l'on peut associer à beaucoup de choses (schémas sous-jacents): l'image d'une montagne, la forme d'une maison, la notion de hiérarchie. Il pourrait même représenter la classe des triangles.

Les meilleurs index picturaux seraient peut-être ceux qui arriveraient à mi-hauteur d'une échelle imaginaire d'abstraction ayant à une extrémité des reproductions réalistes et à l'autre des images très abstraites. Les pictogrammes utilisés lors des Jeux olympiques de Montréal constituent un bel exemple de ce mélange d'abstrait et d'iconique. (voir fig. 1.3). Ces images sont visuellement frappantes et facilement identifiables (dans un certain contexte) sans pourtant chercher à reproduire mécaniquement les aspects apparents des entités qu'elles dépeignent. En fait, ces images ont été composées à partir d'ensembles de formes géométriques simples: cercles, carrés, rectangles, et autres. Elles sont des indicateurs puissants non seulement en raison de la clarté du contexte mais aussi parce qu'elles remplissent deux fonctions

simultanément. Elles véhiculent, d'une part, les qualités iconiques des objets rattachés à un événement sportif donné et semblent incarner, d'autre part, les patrons sous-jacents plus abstraits de forces -- intensité, tension, mouvement, équilibre -- qui font partie de la conception mentale que nous avons d'un sport donné. Ces qualités n'ont aucun corrélatif physique mesurable dans le monde réel. Par exemple, l'image à l'extrême gauche de la figure 1.3 semble avoir mieux capté la "dynamique visuelle" d'un joueur de soccer en train de botter le ballon mieux qu'une photographie d'un vrai joueur de soccer n'aurait pu le faire. La simplicité et le caractère abstrait des éléments graphiques nous aident à percevoir les forces génériques du mouvement, de l'équilibre, de la force, etc., qui sont inhérentes à l'événement; ces éléments graphiques représentent néanmoins "l'image iconique" d'un joueur de soccer. Curieusement, on constaterait sans doute qu'un index iconique efficace possède certaines des propriétés qui caractérisent toute oeuvre d'art. Il joue simultanément le rôle d'icône et de symbole. Autrement dit, un bon index pictural peut être l'image d'un objet ou événement reconnaissable et véhiculer en même temps le patron "plus profond" des forces génériques qui y sont sous-jacentes.

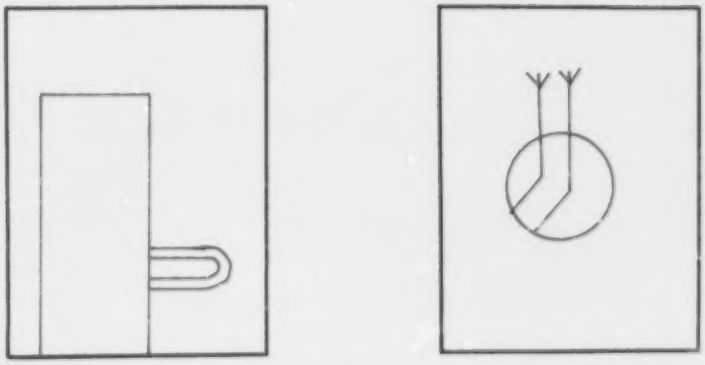
1.5. L'importance du contexte

On pourrait, pour résumer en un seul message tout ce que nous avons exposé jusqu'à maintenant, dire que ce serait une erreur d'utiliser des images comme substituts autonomes d'énoncés verbaux. Mais cela ne ferait que lancer un débat stérile sur la valeur comparée des mots et des images et nous détourner de la véritable question, qui est l'importance du contexte dans la compréhension. Il s'agit notamment de savoir comment les mots peuvent créer un contexte pour faciliter la compréhension des images et comment celles-ci peuvent à leur tour contribuer à la compréhension linguistique. Depuis quelques années, les psychologues de la cognition se consacrent de plus en plus à l'étude du rôle du contexte dans la compréhension. Beaucoup d'expériences

démontrent que, sans le contexte approprié, un message linguistique ou pictural peut-être difficile à comprendre et que l'on peut moins facilement le recevoir ou s'en souvenir. Aux fins de la présente étude, nous nous sommes surtout intéressés à des expériences qui montrent l'efficacité contextuelle réciproque des modes symboliques du langage et des images.

1.5.1. Qu'apportent les mots aux images?

Examinons d'abord comment le langage peut servir de contexte à la compréhension des images. Dans une expérience importante, Carmichael, Hogan et Walter (1932) ont montré comment un énoncé linguistique peut amener les gens à reproduire différemment de simples croquis (par exemples des cercles qui se recoupent) présentés par brefs intervalles à l'aide d'un tachistoscope. Par le simple fait de voir le nom d'un des deux objets possibles (lunettes et haltères) avant l'exposition de l'image, les sujets sont portés à reproduire un dessin qui ressemble à l'objet nommé. Plus récemment, Bower, Karlin et Dueck (1975) ont fait une démonstration intéressante du pouvoir du langage comme contexte dans la compréhension picturale. Ils ont demandé aux participants de comprendre et de mémoriser des "droodles" (dessins-attrapes) (voir fig. 1.4). Ces dessins sont difficiles à comprendre et à mémoriser car ils montrent des fragments insolites d'objets que l'on arrive difficilement à identifier sans indices. Les sujets à qui l'on a présenté des droodles accompagnés d'indices verbaux ont beaucoup mieux réussi les épreuves de mémorisation que les sujets témoins à qui l'on n'a pas donné de descriptions verbales. Les auteurs expliquent le meilleur rendement des premiers sujets par le fait que la description fournit un "schéma" ou un "cadre" -- c'est-à-dire un contexte qui permet d'assimiler l'image dans une structure logique de connaissances existantes. En d'autres termes, et en rapport avec nos réflexions précédentes, la phrase donne une bonne "théorie" de l'image et facilite ainsi la résolution de problèmes du processus interprétatif.



a b

Figure 1.4. "Droodles" (dessins-attrapes): a) Un nain jouant du trombone dans une cabine téléphonique;
b) Une autruche qui s'est aperçue trop tard que c'était du sable mouvant.
(Traduction libre)

Source: Bower, Karlin et Dueck (1975)

Ces expériences, qui montrent comment les phrases activent les structures cognitives qui sont peut-être nécessaires à la compréhension des images, constituent des cas extrêmes d'une fonction très importante du langage dans la communication de tous les jours. Cette fonction encore méconnue des psychologues jouera probablement un rôle important dans le domaine de Télidon. Il s'agit de la capacité que possède une description verbale d'orienter le raisonnement qui découle de la perception visuelle. En d'autres termes, une fois qu'on s'est fait une représentation mentale d'un événement perceptuel ou pictural, le fait d'écouter ou de lire une description peut transformer cette représentation mentale initiale de façon importante -- c'est-à-dire l'influencer afin de révéler des sens spécifiques et parfois inhabituels. Voici une brève énumération de certains processus qu'une description verbale peut déclencher en rapport avec une image: diriger l'observateur afin qu'il sélectionne un seul élément parmi plusieurs éléments d'une image ("regardez la maison à toit de tuiles rouges"); mettre en relief certaines caractéristiques communes ("remarquez comme les deux chambres à coucher sur ce plan ont la même forme de losange"); attirer l'attention sur des relations spatiales ("le niveau d'eau dans la cruche A est plus élevé que celui de la cruche B"); donner des explications aux actions dépeintes ("l'homme au chapeau noir sourit car il vient de gagner à la loterie"); aider l'observateur à percevoir des relations structurales de forme dans l'art ("les deux figures enlacées forment un bloc rectangulaire"). Le but de ces exemples est de nous rappeler qu'une bonne part de la communication quotidienne exige que nous comprenions de quelle façon les descriptions verbales des autres se rapportent aux scènes que nous observons. Il en sera vraisemblablement de même dans le cas de Télidon; l'utilisateur, au lieu d'interpréter une phrase ou une image isolément, sera souvent appelé à intégrer l'information provenant de sources picturales et linguistiques.

Les mots peuvent parfois exercer sur l'interprétation des images une influence très marquée dont les répercussions seront très concrètes. Loftus et Palmer (1974) ont montré à un groupe de personnes un film où deux voitures entraient en collision. Après le visionnement, on a demandé à une partie du groupe d'évaluer à quelle vitesse allait la première voiture lorsqu'elle a "heurté" la seconde. On a ensuite demandé aux autres d'évaluer à quelle vitesse allait la première voiture lorsqu'elle a "embouti" la seconde. Quelque temps après, on a posé aux deux groupes une série de questions sur l'accident dont celle-ci: "Avez-vous vu des éclats de verre sur la chaussée après l'accident?" Ceux à qui on avait posé la question contenant le mot "embouti" ont été plus nombreux à répondre qu'ils avaient vu des éclats de verre que ceux qui avait entendu le mot "heurté". Cette expérience à "question dirigée" démontre de façon convaincante comment les mots peuvent activer des structures cognitives complexes (schémas) qui agissent sur les représentations mentales d'événements perceptuels.

1.5.2. Qu'apportent les images aux mots?

Des expériences menées récemment sur la cognition nous révèlent que non seulement le langage peut guider notre compréhension d'événements picturaux, mais aussi qu'une image peut servir de contexte pour faciliter la compréhension verbale. Une expérience bien connue faite par Bransford et Johnson (1972) le démontre éloquemment :

> Si les ballons éclataient, le son ne pourrait plus être entendu car tout l'attirail se trouverait trop loin de l'étage. Si la fenêtre était fermée, elle empêcherait également le son d'être transmis étant donné que la plupart des immeubles sont généralement bien isolés. Puisque toute la manoeuvre dépend du flux constant du courant électrique, un bris du fil pourrait également causer des problèmes. Bien sûr, notre homme pourrait toujours crier mais la voix humaine n'est pas assez puissante pour porter si loin. Autre problème, les cordes de l'instrument pourraient casser. Il n'y aurait alors aucun accompagnement au message. Il est évident que la situation idéale comporterait moins de distance. Il y aurait alors moins de complications possibles. C'est la communication face à face qui risquerait le moins de rencontrer des obstacles.

Les sujets à qui l'on a fait entendre ce passage sans aucun autre contexte l'ont trouvé très difficile à comprendre et ont effectivement obtenu de piètres résultats lors d'épreuves de rappel des idées maîtresses du passage. Regardez maintenant le dessin de la figure 1.5. Les sujets à qui l'on a montré ce dessin avant de leur faire écouter le passage l'ont trouvé facile à comprendre et ont mieux répondu lors de tests de rappel. La différence entre les deux groupes est, bien sûr, attribuable à l'absence de l'image dans le cas du premier groupe; en effet, l'image contient des objets et des relations qui constituent un "fondement conceptuel" en l'absence duquel les idées sous-jacentes aux phrases sont difficiles à assimiler. Cette expérience est analogue à une étude menée par Bower et alii, dont nous faisons état dans la dernière section pour montrer comment les mots peuvent servir de contexte à la compréhension des images.

Les recherches portant sur le rôle de l'imagerie dans l'apprentissage verbal pourraient aussi nous éclairer quelque peu sur l'apport des images à la compréhension des mots. En fait, le mot "apprentissage" ne convient pas dans ce cas, car le paradigme de recherche prédominant consistait à demander aux sujets d'associer des mots (par exemple, piano, cigare) qui normalement n'ont pas de rapport entre eux. Lorsque, par la suite, on montre au sujet un élément de la paire -- par exemple, piano -- il essaie de se rappeler l'autre élément. En d'autres termes, ces études de "paires associées" ne portent pas vraiment sur l'apprentissage au sens d'utiliser des images pour faciliter l'acquisition de nouveaux concepts en mathématiques ou en sciences, ou encore montrer à quelqu'un comment changer un pneu. Les recherches portant sur le rôle de l'imagerie dans l'association verbale révèlent néanmoins certains faits intéressants, entre autres, pourquoi certaines images arrivent mieux que d'autres à rendre les associations de mots plus faciles à retenir. Mes commentaires sur ces recherches seront très brefs puisque Muter (1980) a fait une excellente analyse du rôle des

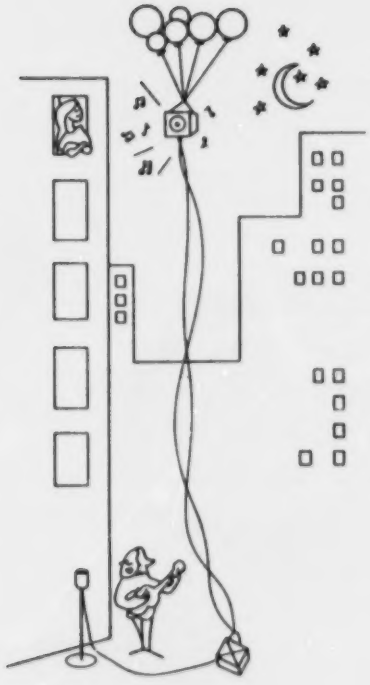


Figure 1.5. Dessin extrait de l'expérience de Bransford et Johnson (1972).

images et de l'imagerie dans le langage verbal et des répercussions de leur utilisation à des fins éducatives sur les systèmes vidéotex.

Imaginez comment vous pourriez associer les mots "piano" et "cigare" de sorte que vous puissiez vous souvenir de l'un des mots si on vous donnait l'autre comme indice à un moment ultérieur. Wollen, Weber et Lowry (1972) ont examiné dans quelle mesure différents types d'images peuvent aider les gens à faire de telles associations entre des paires de mots. Ils ont utilisé quatre types de dessins, dont vous trouverez des exemples à la figure 1.6. (1) non interactif, non bizarre, (2) non interactif, bizarre, (3) interactif, non bizarre, et (4) interactif, bizarre. Ils ont constaté que les paires associées étaient plus faciles à retenir lorsqu'elles étaient composées d'images "interactives" (le caractère bizarre étant sans importance) plutôt que d'images non interactives. Ce qu'il y a d'intéressant dans cette constatation dépend de la signification qu'on attribue à une image "interactive". Malheureusement, les auteurs nous éclairent peu sur le sujet. Si on interprète cette constatation à la lumière des propos que nous avons tenus antérieurement, nous pouvons dire qu'une image "interactive" fournit au sujet un "fondement conceptuel" ou structure cognitive existante qui lui permet d'établir des relations conceptuelles entre des éléments qui sont normalement considérés comme bien distincts. Cette étude révèle que les images constituent des moyens efficaces d'établir de telles relations et qu'elles facilitent par conséquent l'acte associatif.

1.5.3. Interaction des mots et des images

Nous avons montré comment une image peut faciliter la compréhension d'un texte et, inversement, comment le texte peut être indispensable à la compréhension du contenu d'une image. Jusqu'à maintenant, toutefois, nous ne l'avons fait qu'en sens unique, c'est-à-dire du texte à l'image ou de l'image au texte. En terminant la présente section, j'aimerais attirer l'attention sur une autre relation possible, mais encore

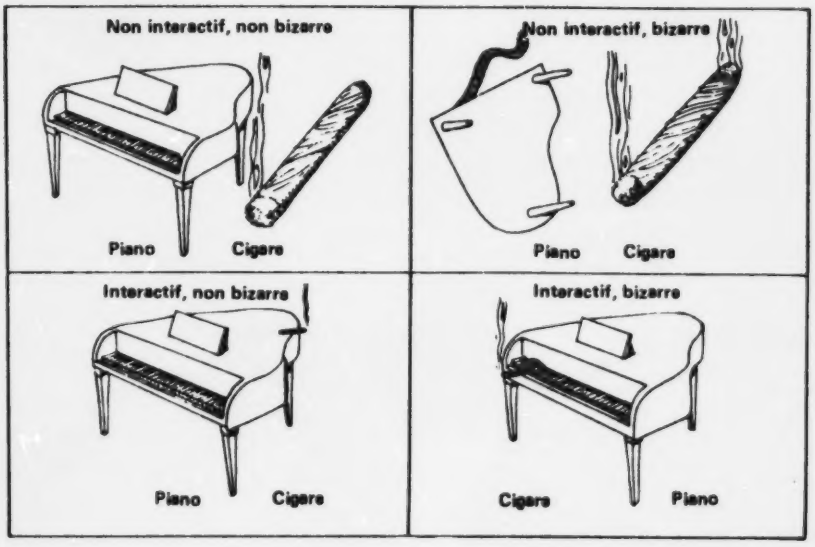


Figure 1.6 Exemples de quatre types de dessins qui accompagnaient les paires de mots présentées dans l'expérience de Wollen et alii. (D'après Wollen, Weber et Lowry, 1972.)

méconnue, entre la signification des mots et celle des images: l'interaction.

Il peut arriver que dans la communication de tous les jours (y compris dans le cas de Télidon) les gens soient appelés à intégrer l'information provenant de scènes visuelles et de descriptions verbales. Il est possible, par exemple, qu'en regardant une image accompagnée d'un texte sur Télidon, on comprenne à la fois le texte et l'image pris séparément, mais que l'on ne saisisse pas immédiatement la relation entre eux. De telles situations peuvent se présenter notamment dans des contextes scientifiques ou pédagogiques où l'association de mots et d'images vise à apporter de nouvelles connaissances. Dans de tels cas, le récepteur aura peut-être à s'engager dans une sorte de "résolution de problèmes" exigeant la modification ou la reformulation des sens sous-jacents aux mots et images, afin de déterminer la relation qui existe entre eux selon le contexte.

Pour bien comprendre comment l'interaction des mots et des images peut en transformer ou contraindre la signification l'un de l'autre dans des contextes bien précis, voyons cet exemple extrait d'une étude de Mills et Bregman (1980). La plupart des gens qui examinent la bande dessinée de la figure 1.7. y voient la représentation d'un événement tel que "un pied écrasant un insecte", où un point noir "écrasé par le mouvement descendant du triangle". Mais, si l'on vous demande de voir comment ce dessin peut être une analogie de l'événement décrit par une autre phrase "une voiture entrant en collision avec un poteau", il semble y avoir peu de rapport entre cette phrase et le type d'événement représenté par le dessin. Après une brève hésitation, toutefois, la plupart des gens voient comment le dessin peut correspondre à la description de l'événement "une voiture entrant en collision avec un poteau". Qu'est-ce qui a changé pour rendre l'analogie possible? Ordinairement, le changement intervient au niveau des notions implicites

Figure 1.7. Un seul dessin qui représente deux événements distincts: (a) Un pied écrasant un insecte (vue de profil); (b) Une voiture entrant en collision avec un poteau (vue en plongée).

Source: Mills et Bregman (1980)

que l'observateur du "point de vue" remarque. Si le dessin représente une vue "en plongée" ou "à vol d'oiseau" -- au lieu d'une vue de profil comme dans le cas de la phrase "un pied écrasant un insecte" -- les éléments visuels concordent avec la signification de la description verbale. Ce qu'il faut remarquer, toutefois, c'est que le passage d'une vue "de profil" à une vue "en plongée" modifie également les significations qui avaient été attribuées à la fois aux mots et aux images. Prenons, par exemple, le point noir du dessin. Au départ, le sens attribué à cet élément graphique peut se limiter à quelque chose comme "objet rond, noir", où le point de vue n'est nullement précisé ou alors on suppose une vue de profil. Afin d'introduire la notion de "poteau", cette signification attribuée au point noir dans notre modèle mental de l'événement doit être modifiée ou révisée comme suit: "point noir représentant un objet haut, droit et cylindrique vu en plongée". Mais remarquez qu'il n'y a pas que le sens attribué aux éléments graphiques de l'image qui change. Le sens nucléaire du mot "poteau" doit également subir certaines modifications selon le contexte de l'image. Si nous supposons qu'une partie du sens archétypal du concept "poteau" encode la connaissance que nous avons de l'apparence archétypale des poteaux (haut, droit, cylindrique et vu de profil), on doit modifier certaines de ces suppositions afin de les faire concorder avec la structure visuelle du point noir du dessin. Le mot "poteau" dans ce contexte doit se rapporter à un poteau vu en plongée afin de concorder avec la forme du point dans le dessin. L'important est de voir comment les sens initiaux, nucléaires, que nous attribuons aux images et aux mots peuvent interagir ou s'infléchir mutuellement (par exemple, subir des transformations) afin de répondre aux exigences de l'acte interprétatif spécifique.

Parce que les psychologues et les linguistes se sont plutôt consacrés jusqu'à maintenant à l'étude du "langage" et de la "perception" en tant que sous-systèmes distincts dans la pensée, on sait très peu de choses sur les types d'interaction, comme dans l'exemple

ci-dessus, qui peuvent se produire entre les significations des mots et des images. C'est malheureux, car nous croyons que de telles interactions se trouvent au coeur même de beaucoup de formes d'acquisition du savoir dans les arts, les sciences et peut-être aussi dans la vie de tous les jours. En résumé, nous avons essayé de montrer dans la présente section que les rapports entre le texte et les images peuvent être assez complexes. Il peut arriver que l'utilisation d'une image sur Télidon ne serve pas tout simplement à illustrer un texte, ou encore qu'un texte ne se limite pas tout simplement à décrire ce qui est visible sur l'image. Il peut y avoir des situations -- notamment dans des contextes pédagogiques -- où l'utilisateur est appelé à entreprendre une sorte de "résolution de problèmes" afin de voir comment une description verbale s'applique à un événement pictural.

1.6. Résumé et conclusions

Le système Télidon peut fournir des renseignements à l'utilisateur sous forme de textes et sous forme d'images de toutes sortes. Cela nous amène inévitablement à nous poser des questions comme celles-ci: "Quel rôle joue l'imagerie visuelle dans la conception des pages d'information dessinées sur Télidon? "S'il faut choisir entre une description verbale et une image pour communiquer une idée, lequel est le meilleur?"; "Puis-je mieux l'exprimer en images qu'en mots?". Ces questions, bien qu'elles aient une signification pratique pour Télidon, nous mènent tout droit à des questions théoriques complexes qui relèvent de la perception et de la cognition humaines, notamment en ce qui concerne la nature du langage, de la perception et de la représentation picturale. Malheureusement, les théories de la communication symbolique ne sont pas suffisamment avancées pour que l'on puisse répondre d'emblée à des questions telles que "lequel est le meilleur?". Lorsqu'il s'agit de déterminer "le meilleur" il faut tenir compte des caractéristiques propres à l'utilisateur, de la nature du travail et d'autres facteurs. Toutefois, notre analyse sélective des théories de la cognition et de la

perception nous a permis d'élucider un certain nombre de questions générales concernant la nature des images et les rapports entre celles-ci et les mots, ce qui peut être utile à une réflexion sur le rôle des médias linguistiques et picturaux sur Télidon. Voici un résumé de ces questions:

(1) Il faut se méfier de l'argument très répandu selon lequel l'image visuelle devrait remplacer le langage comme principal moyen de communication sur Télidon -- du moins en tant que "média autonome". Le langage naturel est le média par excellence pour formuler des propositions destinées à d'autres personnes. Une image peut également véhiculer des "propositions", mais le problème est que, en l'absence d'un certain contexte contraignant (habituellement une description verbale), nous ne savons pas exactement quelle proposition elle veut suggérer parmi le nombre infini de celles qu'elle peut suggérer. À mon avis, de telles ambiguïtés sémantiques rendent peu probable l'avènement d'un langage "purement" pictural. De plus, les codes picturaux auraient fort à faire pour égaler la capacité du langage qui nous permet de formuler des propositions sur des événements passés, futurs et conditionnels ainsi que des chaînes d'inférences logiques. Encore une fois, on ne nie pas que les images soient utiles ou même meilleures que les mots dans certains cas. On essaie plutôt de montrer qu'on ne peut pas compter uniquement sur les images pour communiquer un message.

(2) Nous avons vu que le problème de polysémie des images trouve son origine dans la nature même de l'appareil conceptuel humain -- instrument souple qui se sert d'un ensemble fini de structures cognitives générales ou "schémas" pour donner des "descriptions" définies de données spécifiques. En ce qui

concerne la conception visuelle, nous avons vu qu'une même donnée visuelle peut engendrer nombre de descriptions différentes à partir de différentes combinaisons de schémas -- souvent à la suite de changements dans le contexte. Cette caractéristique confère une grande souplesse au système conceptuel humain, mais nous permet également de mieux comprendre pourquoi les images posent des problèmes dans certains domaines d'application qui intéressent Télidon, telle la conception de schémas de classification en vue de l'extraction automatique d'information picturale par ordinateur; cette même souplesse peut poser des problèmes d'utilisation si les "index iconiques" sont mal conçus. Nous avons examiné certaines des caractéristiques indispensables à un bon index iconique et avons avancé qu'un bon index iconique fonctionnerait simultanément comme "symbole" et comme "icône" et, par conséquent, posséderait certaines des caractéristiques que l'on retrouve dans toute oeuvre d'art. Cela signifie qu'un index iconique efficace doit, d'une part, utiliser des éléments graphiques abstraits pour faire ressortir des qualités génériques et, d'autre part, se présenter sous la forme reconnaissable d'une "icône" -- c'est-à-dire une représentation visuelle d'une sorte d'objet, de situation ou d'événement bien précis. Nous avons vu en outre que, d'après les recherches effectuées sur la catégorisation humaine, y compris les recherches portant sur la nature du langage gestuel (ASL), certains objets sont plus "représentatifs" que d'autres de même catégorie et que, par conséquent, ils seraient plus efficaces comme index picturaux.

(3) Enfin, nous avons cherché à démontrer qu'au lieu de substituer des images aux mots ou inversement, il faut plutôt voir comment mots et images se complètent pour diriger le récepteur vers le sens précis d'un message. Nous avons cité certains

travaux de recherche qui montrent que: (1) les images peuvent être ambiguës ou dénuées de sens si elles ne sont pas accompagnées d'un contexte verbal approprié; (2) un passage verbal peut être ambigu ou difficile à comprendre à moins d'être accompagné d'une image pertinente; (3) la juxtaposition des mots et des images produit parfois des effets réciproques qui modifient la signification première des uns et des autres -- notamment dans les contextes pédagogiques. En d'autres termes, il peut y avoir des <u>interactions</u> entre la signification des mots et celle des images.

CHAPITRE II
LIRE LES IMAGES: LE RÉALISME EST-IL CE QU'IL Y A DE MIEUX?

Ce chapitre est axé sur ce qui sera probablement la fonction prédominante de l'imagerie Télidon, surtout au début: la représentation picturale, c'est-à-dire l'emploi d'une image pour représenter une chose à un niveau d'abstraction inférieur à l'image elle-même (voir la section 1.4.1). Par représentation picturale, nous entendons simplement l'emploi d'une image pour renseigner un observateur sur l'apparence d'une chose, qu'elle lui soit familière (un arbre) ou non (un nouvel outil). Nous nous proposons donc d'examiner quelques-unes des idées qui ont cours concernant la façon dont nous reconnaissons les objets représentés par des images, particulièrement par des dessins au trait et des caricatures, et de montrer que ces idées pourraient remettre en question l'opportunité du réalisme photographique sur Télidon. Plus précisément, nous étudierons comment nous pourrions exploiter l'efficacité avec laquelle l'intelligence humaine comprend les dessins schématiques et les caricatures de façon à optimiser Télidon en tant que moyen imagier, c'est-à-dire produire des dessins simples et des graphiques schématiques.

2.0 Apprenons-nous à lire les images?

Il est possible qu'une image ne puisse pas, en soi, rendre superflues les descriptions verbales, mais elle peut fournir des renseignements utiles, dont les plus élémentaires concerneront, entre autres, l'apparence, la forme, la texture et les relations spatiales des objets. À première vue, ceci n'a rien de mystérieux ou de remarquable. Une image peut représenter un objet réel, par exemple, une pomme, parce qu'elle ressemble, dans un certain sens, à une pomme. La façon traditionnelle d'envisager ce processus de ressemblance consiste à voir l'image comme étant une surface traitée artificiellement qui transmet à l'oeil le même faisceau de rayons lumineux qu'une scène réelle que

regarderait directement un observateur. En fait, c'est ainsi que J.J. Gibson (1954) explique la perception de l'image. D'après Gibson, l'image est une surface pigmentée qui, à cause de sa façon de réfléter la lumière, se substitue à des objets réels qui sont dans leur milieu naturel.

L'explication de Gibson est simple et montre pourquoi nous ne ressentons pas le besoin d'apprendre à lire les images comme c'est le cas pour la langue écrite. Toute personne qui a appris à voir peut voir une image. Pourtant, malgré la simple logique de cette définition, certains critiques, notamment le philosophe Nelson Goodman, (1968, 1971), se sont élevés contre la théorie de la ressemblance de la représentation. D'après Goodman, la notion de la ressemblance n'explique pas pourquoi nous pouvons comprendre les images. Il est d'avis que notre compréhension de l'image s'apparente davantage au processus d'apprentissage d'une langue dans sa dépendance des coutumes, des conventions et de la culture. Il y a, dans un sens très large, un rapport entre Télidon et le débat entourant la nature d'une image parce que ce débat fait intervenir la question de l'universalité des images en tant que moyen de communication. Une image ou un graphique présenté sur Télidon sera-t-il immédiatement et universellement compris par tous les voyants? L'apprentissage de la lecture des images ne serait-il pas plutôt un processus culturel nécessitant une période d'adaptation ou une formation spécialisée? On doit, de toute évidence, posséder une formation spécialisée pour lire certaines images, par exemple les radiographies, les cartes topographiques et les schémas de circuits et nous examinerons quelques images de ce genre dans le chapitre III. Ce qui nous intéresse ici, c'est la lecture d'images réalistes. Qu'est-ce qui nous permet de déterminer que des signes tracés sur une feuille de papier représentent une table, un chandail et ainsi de suite?

Goodman et d'autres, dont Gombrich (1960), ont souligné l'existence de différences marquées entre la perception directe et la perception des

images. Par exemple, il y a deux importantes sources d'information qu'une image unique ne peut offrir: la disparité rétinienne (l'impression de profondeur due au fait que, lorsqu'une personne regarde une scène, ses deux yeux reçoivent des projections partiellement chevauchantes) et la parallaxe monoculaire (les positions relatives des projections des objets sur la rétine de l'observateur se déplacent par rapport à lui lorsqu'il bouge la tête, ce qui fait apparaître de nouvelles parties de la scène et en fait disparaître d'autres). De toute évidence, ni la disparité rétinienne ni la parallaxe monoculaire n'interviennent dans l'observation d'une image ou d'une peinture. Par exemple, lorsqu'une personne passe devant une peinture dans une galerie d'art, les objets représentés conservent leurs positions relatives même lorsque la personne change son angle de visée par rapport à l'image. Un autre argument à l'encontre des théories qui assimilent la perception des images à la perception directe concerne la perception de l'exactitude projective de l'image. On a signalé par exemple que les dessins d'objets solides ne respectant pas strictement les règles de la perspective linéaire, par exemple un cube dont la surface titrée comporte des lignes parallèles au lieu de lignes convergentes ou un disque qui est circulaire plutôt qu'en forme d'ellipse comme il le serait dans l'image rétinienne (voir la figure 2.0), ne troublent pas les observateurs. Ce sont des arguments de ce genre qu'ont avancés Goodman et d'autres pour démontrer qu'il est impossible d'assimiler la lecture d'une image à la perception directe et que la compréhension des images dépend plutôt des coutumes, des conventions et de l'expérience antérieure d'une personne et ressemble davantage à l'apprentissage de la compréhension des symboles d'une langue.

Dans une étude des différentes recherches psychologiques portant sur la question, Perkins (1979) conclut que la vérité concernant la représentation par images se situe très probablement entre ces deux extrêmes. On possède des preuves à l'appui de chacune des théories.

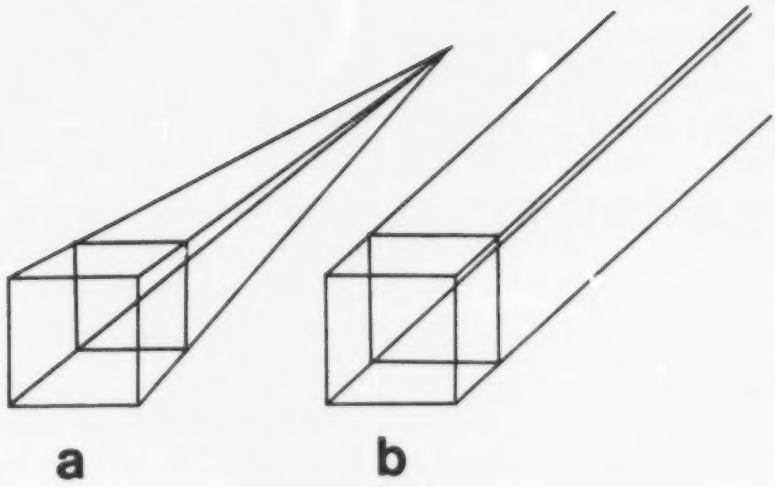


Figure 2.0. Le dessin (a) peut représenter un cube: les lignes convergentes respectent les lois de la perspective linéaire. Le dessin (b) ne peut pas, à proprement parler, représenter un cube, car les lignes qui en prolongent les côtés sont parallèles.

En général, elles découlent d'expériences portant sur de jeunes enfants ou d'études interculturelles de populations indigènes ne connaissant pas la représentation par images.

L'étude de Hochberg et Brooks (1963) dont les sujets sont de jeunes enfants est parmi les plus connues. Cette étude renforce l'argument selon lequel la lecture des images n'exige pas de processus d'apprentissage distinct. Dans une étonnante démonstration de dévouement à la science, Hochberg et sa femme ont élevé leur enfant jusqu'à l'âge de dix-neuf mois sans le laisser voir d'images. Aucun périodique ou livre d'images n'entrait dans la maison. Les étiquettes étaient enlevées des boîtes de conserves. Lorsque l'enfant eut deux ans, Hochberg le soumit à une épreuve. Il lui montra une série de dessins-contours simples, de dessins élaborés et de photographies d'objets d'usage courant dont l'enfant avait déjà appris le nom. Ce dernier reconnut sans difficulté les objets qu'on lui montrait, quel que soit le mode de représentation employé. L'opinion contraire, c'est-à-dire que les enfants doivent apprendre à lire les images, tire ses appuis des recherches de Bower (1964 et 1966) qui démontre que de jeunes enfants conditionnés à répondre à la présence d'objets réels ne transféreraient pas la réponse à des images représentant ces objets.

Les études interculturelles portant sur la compréhension de l'image ont également donné des résultats partagés, mais corroborent, en général, ce que Perkins a nommé la théorie de "l'objet réel". Après une brève période d'adaptation, les personnes appartenant à des cultures n'étant pas familières avec les images reconnaissent sans grande difficulté des objets, et la plupart des relations spatiales, représentés dans des photographies ou des dessins (voir Hagen et Jones, 1978, et Jahoda et McGurk, 1974). Perkins cite une exception intéressante à ce qui précède. Duncan, Gourlay et Hudson (1973) ont mené une expérience portant sur la compréhension des instruments graphiques servant à

exprimer le mouvement. Par exemple, des Africains ruraux, contrairement aux Sud-Africains de race blanche, ont rarement interprété comme étant une représentation du mouvement une série de têtes disposées dans différentes positions sur une même page et des têtes accompagnées de traits indiquant une ligne de fuite.

Un point sur lequel il semble que les différences entre la perception d'une image et la perception de l'objet réel ne varient guère concerne l'aptitude des gens à porter des jugements sur les propriétés spatiales avec une exactitude métrique: par exemple, Olson, Pearl, Mayfield et Millar (1976) ont examiné l'aptitude d'adultes et d'enfants à juger la pente de boîtes représentées par des images. Ils ont remarqué que les adultes et les enfants évaluent de façon assez juste l'orientation de boîtes réelles, mais que les deux groupes, lorsqu'il s'agit de déterminer l'inclinaison réelle de boîtes représentées par des images, commettent des erreurs qui atteignent parfois 20 degrés. D'après Perkins, une telle régression vers le plan avant dans l'observation d'images serait un compromis que fait le système visuel en vue de concilier la planéité déclarée de l'image et les renseignements fournis par l'image sur la profondeur. Lorsqu'on minimise la planéité de l'image, l'appréciation est plus juste (Attneave, 1972).

Ces résultats nous permettent-ils d'établir des généralisations sur la perception des images? Les images forment-elles un langage circonscrit par des coutumes et des conventions ou les comprend-on parce qu'elles transmettent des renseignements aux yeux à la façon des objets réels? Les travaux de Perkins nous permettent de résumer la situation de la façon suivante. Fondamentalement, les preuves corroborent une théorie dite "de l'objet réel" de la compréhension des images, mais avec d'importantes réserves. On comprend les images parce qu'elles présentent les renseignements à peu près de la même façon que les scènes réelles. D'autre part, on peut prévoir que des personnes de groupes d'âge et de groupes culturels différents comprendront des images réalistes sans passer par une longue période d'apprentissage méthodique

comme celle dont on a besoin pour apprendre une langue. Cependant, le fait que certains moyens de représentation exigent un apprentissage (les lignes de mouvement, par exemple) et la persistance de certains types d'erreurs (par exemple, dans l'appréciation des inclinaisons) donnent à penser que les coutumes et les conventions jouent un rôle dans certains aspects de la compréhension des images. Nous devons cependant insister, à l'instar de Perkins, sur la nécessité de ne pas interpréter le rôle que jouent les conventions dans la lecture des images comme étant arbitraires. En effet, alors que la relation entre le mot "pomme" et l'objet réel peut avoir été établie par l'usage (le mot orange aurait pu convenir tout aussi bien), l'utilisation de traits pour représenter des lignes de fuite n'est pas arbitraire dans le même sens. (Par exemple, une ligne pointillée ne représenterait pas aussi bien le mouvement.) Les lignes de fuite rendent plutôt, dans un graphique, ce qui serait un équivalent structurel d'un certain genre de constante perceptive: la trace de lignes que laisserait un objet lorsqu'il passe rapidement devant nos yeux.

Voici comment Perkins résume la situation. Une théorie dite d'objet réel ou "de ressemblance" de la représentation est une approximation de premier ordre de la façon dont les êtres humains comprennent les images; une théorie de deuxième ordre devra tenir compte de l'influence des coutumes, des conventions et de l'apprentissage.

2.1 Le langage des lignes

Une image est donc un instrument remarquable, de compréhension universelle, conçu pour communiquer des renseignements sur l'apparence des choses. (La signification particulière d'une image dépend néanmoins toujours du contexte. Voir le chapitre 1.) Une image est une surface traitée qui nous renseigne sur le monde visuel selon une présentation qui ressemble à peu près à la façon dont le monde visuel les présente. Pour fournir des renseignements utiles sur les apparences, il n'est

cependant pas nécessaire qu'une image reproduise une scène de façon mécaniquement fidèle comme dans une photographie ou une peinture réaliste. Une des plus puissantes méthodes de représentation, vieille comme le monde et pratiquée par toutes les cultures et qui n'enregistre pas point par point l'éclairage de l'objet visible à l'observateur, est le dessin-contour simple.

John Kennedy (1974), un des anciens élèves de Gibson, a étudié la nature des dessins en profondeur. Il a utilisé l'image reproduite à la figure 2.1 pour démontrer qu'on peut rendre par des segments de lignes beaucoup des éléments fondamentaux de la structure optique. Par exemple:

1) une borne occluante (une surface courbe convexe dont une partie est cachée à l'observateur) et une surface d'arrière-plan peuvent représenter le sommet d'une colline et la surface d'une autre colline derrière;

2) une borne occluante sans surface d'arrière-plan, le sommet d'une colline au-dessus duquel le ciel est visible;

3) une arête occluante sans surface d'arrière-plan, le sommet du toit d'une maison au-dessus duquel le ciel est visible;

4) une arête occluante et une surface d'arrière-plan, la fin d'un mur dont le reste de la surface est masqué par la surface qui se trouve derrière et la surface du sol visible au-delà de la fin du mur;

5) un angle dièdre formant un coin concave, deux surfaces planes et visibles d'un mur formant un angle de moins de 180° mesuré dans l'espace contenu par le mur;

6) un angle dièdre formant un coin convexe, deux surfaces d'une maison formant un angle de plus de 180° mesuré dans l'espace entourant les surfaces.

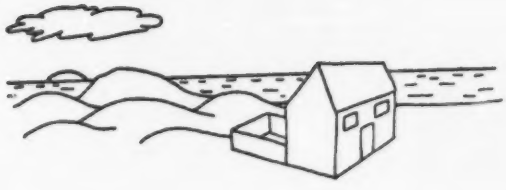


Figure 2.1. Les significations d'une ligne.
Source: Kennedy (1974)

Outre ces caractéristiques de l'aménagement d'une surface, des lignes peuvent représenter des fissures, des ombres, la texture et des points saillants. Les seules caractéristiques du monde visuel que les lignes ne peuvent représenter sont les couleurs isolées et l'espace uniforme.

Comment un simple tracé peut-il représenter autant de caractéristiques de l'environnement visuel? Après tout, il n'y a pas de lignes dans l'environnement naturel, seulement des arêtes, des surfaces et d'autres éléments de ce genre. La capacité qu'ont les dessins de représenter ces éléments ne signifie-t-elle pas que c'est par expérience et par la connaissance des conventions que nous apprenons à les lire? Kennedy réfuterait rapidement cet argument en disant que les lignes sont utiles parce qu'elles rendent, dans un tableau visuel figé, les invariants de la structure optique, particulièrement les discontinuités de la brillance causées par le reflet de la lumière sur les arêtes, les changements de texture, l'agencement de la surface et ainsi de suite. Pour Gibson, ces invariants sont d'une importance capitale dans la perception du monde visuel.

Alors que Kennedy centre ses efforts sur l'analyse des différentes discontinuités que peuvent représenter les lignes, Julian Hochberg (1972) a ajouté un élément fascinant à l'histoire de la compréhension des lignes. Hochberg postule au départ que, pour comprendre la nature des dessins, nous devons d'abord envisager la perception de façon plus fonctionnelle et plus globale comme étant une activité dirigée vers un but et échelonnée dans le temps. Pour comprendre ce processus, dit-il, il est essentiel de comprendre comment le système visuel peut reconstituer une scène à partir de regards discrets successifs.

> Parce que nos yeux n'enregistrent les fins détails qu'à l'intérieur d'une très petite zone fovéale du champ visuel, il nous faut jeter une série de regards dans différentes directions pour nous renseigner sur le monde visuel. Nous procédons par mouvements oculaires saccadés dont les points terminaux sont déterminés d'avance (c'est-à-dire que les saccades sont des mouvements balistiques). Le contenu de chaque regard est donc toujours, dans un certains sens, une réponse à une question au sujet de ce que l'observateur verra si une partie déterminée de la scène vue à la périphérie est amenée à la fovéa. Lorsqu'il regarde un monde normal, l'observateur possède deux sources d'attentes: (i) il est déjà partiellement renseigné sur les formes qu'il doit s'attendre à y voir ainsi que sur les régularités de ces formes et (ii) le champ périphérique étendu de la rétine, dont l'acuité est médiocre et, par conséquent, ne peut que médiocrement capter les détails, indique néanmoins à l'observateur ce que captera son regard lorsqu'il tournera les yeux vers une zone donnée du champ visuel (Hochberg, 1972, p. 65, traduction libre).

D'après Hochberg, les mouvements oculaires sont guidés par des expectatives ou des hypothèses fondées sur des renseignements que fournit la vision périphérique au sujet des éléments que la fovéa est susceptible de trouver. Quels sont donc les renseignements les plus importants pour des organismes mobiles et actifs? Ce sont des données visuelles du genre de celles que Gibson et Kennedy ont si bien décrites, entre autres, les arêtes, les bornes, les gradients de texture et les fissures, des agencements visuels qui signalent à l'organisme de brusques changements dans la configuration du milieu. La sensibilité de la vision périphérique aux discontinuités de ce genre, ordinairement indiquées par des modifications d'éclairage, est logique pour assurer la survie des animaux en mouvement. Mais comment expliquer que des lignes tracées sur du papier puissent faire naître de telles hypothèses concernant des discontinuités importantes? Après tout, comme nous l'avons déjà dit, il n'y a pas de lignes dans le monde; il n'y a que des arêtes, des surfaces et ainsi de suite. D'après Hochberg, si nous comprenons les dessins-contours c'est probablement parce que les structures du système nerveux visuel qui réagissent au contour

délimitant les différentes zones de brillance sont aussi celles qui réagissent aux lignes tracées sur papier. En bref, les lignes tracées sur une feuille de papier indiquent des discontinuités importantes du milieu visuel qui sont du même genre que celles auxquelles "est accordée", si on peut dire, la vision périphérique dans la perception normale. Le système visuel est sensible à ces discontinuités parce qu'elles aident l'organisme à savoir où diriger son regard.

2.2 L'équilibre temps/réalisme: les répercussions de l'utilisation de dessins-contours sur Télidon

Supposons que vous êtes à la recherche d'une maison et que vous n'avez qu'une idée générale du genre de maison que vous cherchez, de la somme d'argent que vous souhaitez consacrer à l'achat, de l'endroit où la maison sera située et ainsi de suite. Votre première tâche sera de vous renseigner sur les maisons qui sont sur le marché afin de circonscrire le champ des possibilités. Si vous n'aviez aucune contrainte de temps, la meilleure façon de procéder consisterait à visiter chacune des propriétés susceptibles de vous intéresser. Une solution de rechange pratique pourrait être d'utiliser un système vidéotex de recherche documentaire, Télidon par exemple, pour faire un survol des possibilités. De cette façon vous pourrez rétrécir le champ à quelques maisons et les visiter par la suite. (Un système idéal vous permettrait d'effectuer la recherche à l'aide de mots-clés pour obtenir la liste des propriétés répondant à vos exigences.) Les images pourraient jouer un rôle utile dans l'extraction des données concernant les diverses propriétés. Examinons la possibilité la plus évidente: l'utilisation d'images pour renseigner l'acheteur éventuel sur l'apparence des maisons.

La première réaction d'un fournisseur d'information serait peut-être de présenter des photographies en couleur en utilisant un degré élevé de résolution. Une bonne photographie de la maison intéresserait sans doute l'utilisateur. Qu'arriverait-il cependant si, pour la visualisation de chaque photo sur son écran de télévision, l'utilisateur devait attendre deux ou trois minutes, par exemple, la reconstitution de l'image ligne par ligne. La construction d'images photographiques sur les systèmes vidéotex actuels entraînera de tels délais si l'on utilise des voies téléphoniques à bande étroite pour la transmission, au terminal de l'utilisateur, d'images de qualité photographique. Chaque point constituant l'image est stocké dans l'ordinateur du centre de traitement selon un code numérique, transmis par lignes téléphoniques et reproduit sur le terminal de l'utilisateur au moyen d'un balayage lent ou en mode de fac-similé. Donc, même si le vidéotex permet la transmission d'images photographiques selon le mode de visualisation par profil de multiplets, il faut compter des délais accrus et des exigences de stockage en mémoire centrale. (On pourrait fort bien restreindre la photographie à une petite partie de l'écran, ce qui diminuerait le délai de transmission tout en réduisant la précision visuelle.

La transmission utilisant ce mode de visualisation soulève une question intéressante, à savoir, quel serait l'effet de tels délais sur l'utilité perçue du système. Si l'utilisateur n'avait que deux ou trois maisons à regarder, une attente de deux ou trois minutes par image ne poserait probablement aucun problème. Par contre, que se produirait-il si, comme ce sera sans doute le cas, l'utilisateur désire obtenir des renseignements sur 15 ou 20 maisons afin d'avoir une idée générale des propriétés disponibles? Des délais de trois minutes par image pourraient avoir des effets cumulatifs intolérables. Dans un pareil contexte, bien que l'on ne connaisse pas encore le seuil de tolérance

par rapport aux délais, même un délai d'une demi-minute pourrait être inacceptable. (Précisons que même si l'exemple utilisé relève du domaine immobilier, les problèmes soulevés par le codage et la transmission de photographies en mode de balayage lent se manifestent tout autant pour d'autres genres d'images.)

Même si les systèmes vidéotex permettent d'extraire des images photographiques réalistes, cette possibilité impose des coûts à l'utilisateur en termes de délais de transmission et des pénalités au fournisseur d'information en termes d'exigences d'entreposage dans l'ordinateur central. En raison de ces coûts, les propos que nous tenions dans la section précédente au sujet de l'utilité des dessins pour la présentation de renseignements sont des plus pertinents, surtout lorsqu'on tient compte de la capacité de Télidon de produire des graphiques linéaires de bonne qualité. Examinons quelques-uns des compromis auxquels le délai de présentation et le réalisme pourraient donner lieu.

Imaginons qu'au lieu de stocker les données nécessaires à la transmission d'images photographiques en couleur de 20 maisons, nous puissions faire des dessins de maisons réelles ou de certains modèles de maisons (voir la figure 2.2) et ne stocker que les données relatives aux lignes. Rien qu'au poste "stockage", nous pourrions réaliser des économies appréciables sur l'ensemble des 20 maisons. Plus important encore, puisque seules les caractéristiques importantes sont mémorisées en termes de simples segments de lignes, on raccourcirait radicalement le temps de visualisation d'une image sur l'écran. Au lieu des minutes qu'exige la transmission d'une photographie, la transmission d'un dessin au trait se compterait en secondes. L'utilisateur réaliserait ainsi de

Figure 2.2. Dessin-contour d'une maison fait à partir d'une photographie en couleur.

grandes économies de temps, ce qui accroîtrait peut-être sa satisfaction à l'égard du système.[1] Les questions d'ordre psychologiques soulevées ici peuvent être reformulées de la façon suivante:

1) Un dessin-contour peut-il représenter une maison, par exemple, aussi bien (ou mieux) qu'une photographie en couleur? D'après les travaux de Kennedy et de Hochberg, la chose n'est pas impossible. Les dessins peuvent fournir des renseignements sur la plupart des sources importantes de la structure optique du monde visuel. En fait, si Hochberg a raison, il serait même plus facile et plus rapide de les extraire d'un dessin que d'une photographie détaillée.

2) Il se peut cependant que cette hypothèse d'équivalence achoppe sur un point important, relié à ce qu'on pourrait appeler la "métaconnaissance" des moyens de communication. Imaginons que nous ayons une photographie et un dessin d'une maison semblable à celle qui est représentée à la figure 2.2. Supposons en outre qu'il soit possible de prouver, au moyen d'un test expérimental, l'équivalence informationnelle du dessin et de la photographie, c'est-à-dire que l'un permet aussi bien que l'autre de répondre à des questions concernant, par exemple, la forme, la dimension ou l'aménagement de la maison. On remarquera

[1] Il y a lieu de signaler qu'on retrouve cette notion de "délai" parmi les facteurs reconnus de l'acceptabilité d'autre matériel de télécommunication; par exemple, la période d'attente entre le moment où l'utilisateur d'un téléphone à clavier signale le dernier chiffre et entend le déclenchement de la tonalité (M. Schaeffler, communication personnelle).

peut-être encore des préférences marquées pour la photographie et ce, pour les raisons suivantes. L'utilisateur peut avoir l'impression qu'une photographie est un enregistrement plus fidèle et plus objectif de l'objet réel, pas nécessairement parce qu'elle ressemble davantage à l'objet réel (une photographie floue de piètre qualité peut être moins exacte, dans un certain sens, qu'un dessin linéaire exécuté avec soin), mais parce qu'il croit que, contrairement à l'artiste qui exécute un croquis, un appareil photographique ne peut mentir. On remarquera toutefois que ce genre de croyance concernant l'objectivité des photographies peut être dénué de tout fondement. L'appareil photo peut fort bien manipuler, déformer et exagérer la réalité, tout photographe publicitaire le sait, tout comme le langage des lignes peut rendre de façon relativement objective certains aspects de l'information.

Ces questions en appellent d'autres, dont l'intérêt n'a d'égal que leur importance par rapport à l'utilisation efficace des images sur Télidon et surtout par rapport à nos connaissances générales de la nature de la communication par les images: comment la confiance qu'inspire aux gens la fidélité apparente de différents types d'images interagit-elle avec la fidélité réelle de ces images et comment cette question de "confiance" interagit-elle avec des variables comme le délai de transmission, la composition des pages et ainsi de suite.

Permettez-moi d'esquisser rapidement quelques-unes de ces interactions et de démontrer comment elles pourraient se prêter à des études empiriques. Il serait utile, pour commencer, de vérifier l'hypothèse selon laquelle le degré d'objectivité que les gens attribuent respectivement aux dessins-contours et aux images photographiques correspond à la capacité informationnelle de ces deux modes de

représentation. On pourrait présenter des exemples des deux genres de représentation et demander aux sujets de les coter d'après des aspects qui ébranleraient l'idée qu'ils se font de l'objectivité des images. La deuxième étape pourrait consister à vérifier la capacité informationnelle des représentations linéaires par rapport à celle des représentations photographiques en soumettant les sujets à une épreuve simple: ainsi, le fait de voir un dessin d'un objet, une maison par exemple, permettrait-il de reconnaître plus facilement cette maison dans un autre médium (en photo, par exemple). On pourrait aussi examiner directement la façon dont un délai de transmission influe sur l'appréciation des modes de représentation linéaire et photographique par les utilisateurs. Il se peut, par exemple, qu'on accepte de sacrifier le réalisme des photographies à des graphiques qui apparaîtraient plus rapidement à l'écran, sous forme de dessin-contour, si l'on avait à utiliser un système vidéotex pour effectuer une recherche documentaire portant sur un vaste éventail d'articles.

Outre le genre d'images et le délai de transmission, une autre variable sur laquelle on pourrait enquêter est la composition des pages. Un seul exemple suffira à illustrer ce point. La graphique des ordinateurs qui permet à un utilisateur de regarder une matrice composée d'images distinctes visualisées simultanément sur son écran vidéo existe.[1] Grâce à une matrice de ce genre, on pourrait exploiter la capacité qu'a l'oeil humain de repérer rapidement, par balayage, un objet cible particulier parmi une grande variété d'objets disposés dans

---

1 Le Machine Architecture Group du MIT a mené, sous la direction de Nicholas Negroponte, des expériences utilisant des affichages d'images multiples. Il a examiné, tout particulièrement, la possibilité d'allier la capacité de mémorisation des vidéodisques à la puissance de traitement des ordinateurs en vue de fournir de nouvelles possibilités aux fins de la manipulation des images (voir le chapitre III, section 3.0.1).

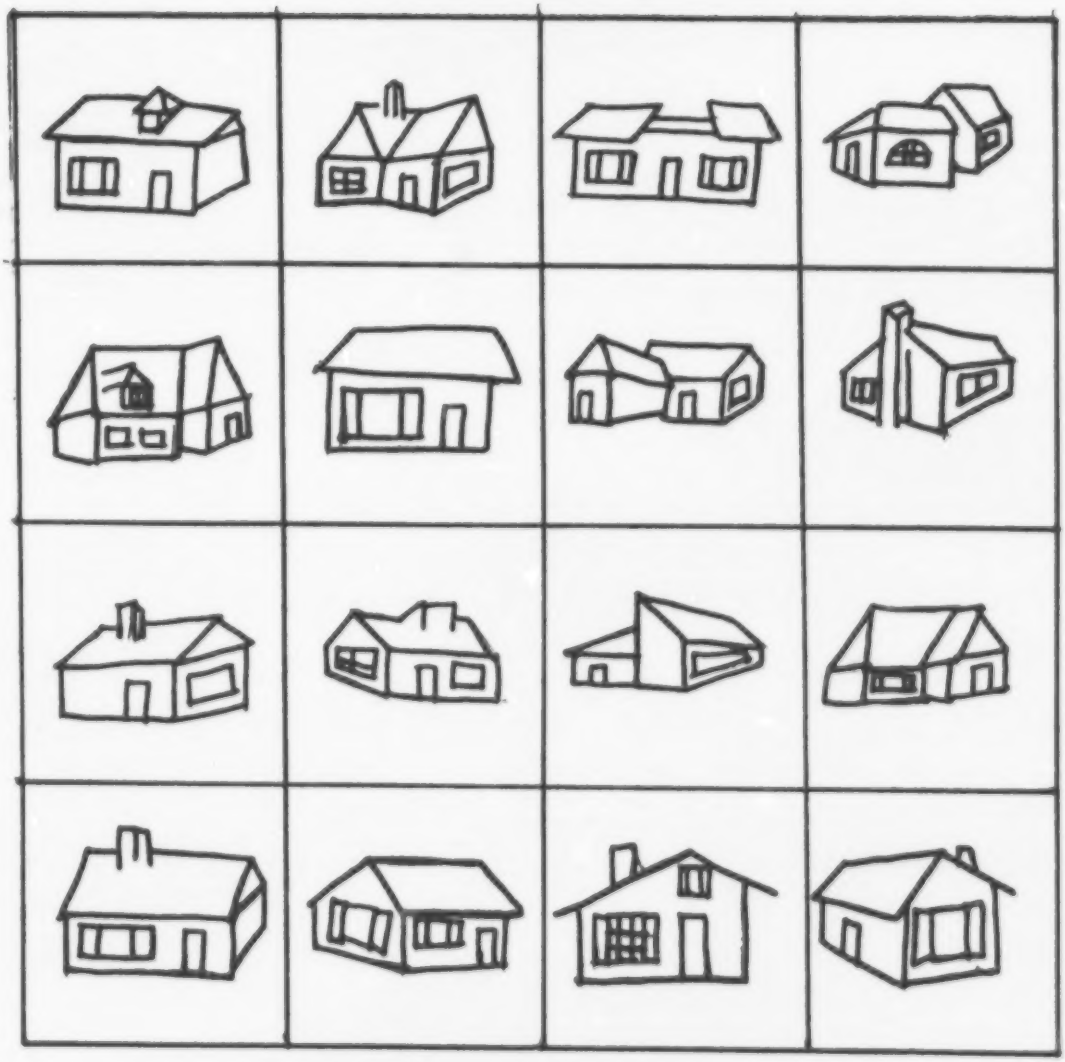


Figure 2.3. Une matrice d'images destinée à un survol de l'information.

un espace donné. De nombreuses applications pourraient découler d'un système de ce genre. Nous pourrions, ici encore, tirer un exemple du domaine de l'immobilier. La figure 2.3 montre des représentations linéaires de 12 maisons, disposées sous forme de matrice.

Des photographies auraient tout aussi bien pu servir de médium de représentation. Un affichage de ce genre pourrait servir d'index imagier. L'utilisateur pourrait parcourir cet index des yeux et y choisir un ou plusieurs articles qui seraient ensuite visualisés sous la forme d'images plus détaillées et de plus grandes dimensions (photographies ou dessins qui pourraient être accompagnés de descriptions verbales). Un index iconique interactif de ce genre serait utile. Mais la question est de savoir s'il est préférable d'employer des photographies ou des dessins linéaires dans une matrice de ce genre. Le délai de transmission sur une voie à bande étroite serait beaucoup plus long dans le cas d'une matrice composée de photographies, de maisons par exemple, que de dessins linéaires. Sur le plan psychologique, cependant, ce qui nous intéresse, en supposant que les délais de transmission soient identiques, c'est de savoir si l'utilisateur saisirait plus facilement les principales caractéristiques des maisons si celles-ci étaient présentées sous forme de dessins plutôt que sous forme de photographies. Pourquoi pourrait-il en être ainsi? D'après Hochberg, c'est parce que les lignes font ressortir les éléments essentiels (les changements soudains) dont le système visuel a besoin pour savoir où regarder ensuite.

2.3 Les photographies, les dessins, les bandes dessinées et les caricatures

Certaines recherches effectuées dans le domaine de l'éducation et de la psychologie ont une incidence directe sur la question du contenu informationnel relatif des dessins et des photographies à haute fidélité. Dwyer (1972, 1978) a découvert par exemple que l'emploi de

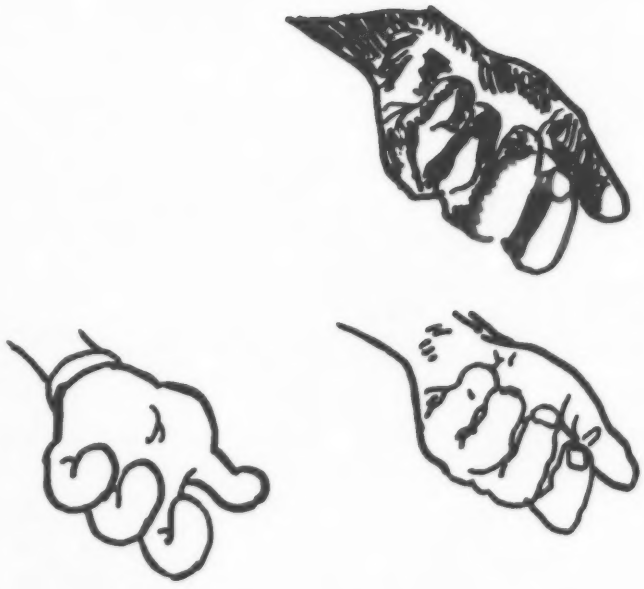


Figure 2.4 Trois des quatre types d'image utilisés dans l'expérience de Ryan et Schwartz. La quatrième image était une photographie.
Source: Ryan et Schwartz (1956)

photographies dans les textes médicaux n'aide pas nécessairement à reconnaître ultérieurement des objets plus que l'emploi de dessins, quoique les couleurs soient utiles dans certains cas. Cependant, Dwyer a aussi découvert que les gens s'attendent à ce que les photographies soient plus révélatrices que les dessins, même si leur rendement ne répond pas à leurs attentes. Comme nous l'avons signalé ci-dessus, cette question de l'efficacité relative que les gens attribuent aux différents modes de représentation mérite d'être étudiée plus à fond dans le contexte de Télidon.

Deux chercheurs, Ryan et Schwartz (1956) ont mené une étude dans laquelle ils font une comparaison directe de l'information véhiculée par plusieurs modes de représentation picturale: photographies à haute fidélité, dessins ombrés détaillés, dessins simplifiés et cartoons. La figure 2.4 montre trois des quatre genres d'images qu'ils ont utilisées au cours de leur expérience. (Il ne manque que les photographies.) On a montré aux sujets des images présentant divers objets (par exemple une main dont les doigts étaient placés dans une position particulière, des lames de couteau orientées dans diverses positions, divers profils d'une machine) selon l'un ou l'autre des quatre modes. Les images ont été projetées de façon tachistoscopique pendant de courtes périodes, les expositions se prolongeant de plus en plus jusqu'à ce que les sujets soient en mesure de reconnaître certaines caractéristiques de l'image, par exemple la position des doigts de la main. Chose intéressante, les sujets ont eu besoin de périodes d'exposition plus courtes dans le cas des bandes dessinées. Les dessins ont exigé des périodes d'exposition plus longues et on n'a constaté aucune différence entre les photographies et les dessins ombrés. Cette étude est importante parce qu'elle soulève des questions non seulement au sujet de la supériorité des lignes sur les photographies, mais, encore plus, au sujet de la nature des bandes dessinées. Pourquoi les sujets ont-ils reconnu plus facilement les bandes dessinées?

2.3.1 Les bandes dessinées concrétisent les schèmes intérieurs

L'efficacité des bandes dessinées et des caricatures est intéressante parce que, d'un point de vue strictement kennedy-gibsonien, ces deux formes de représentation s'écartent d'une des définitions de l'image-contour, soit une surface délimitée par des lignes qui renseigne l'oeil sur une scène de la même façon que le ferait le faisceau de rayons lumineux émanant de la scène elle-même. Les bandes dessinées et les caricatures semblent précisément tirer leur grande efficacité du fait qu'elles ne reproduisent pas exactement ce faisceau de rayons lumineux; non seulement omettent-elles de l'information mais elles transforment et exagèrent l'information par des moyens (encore mal compris) qui semblent faciliter la reconnaissance. L'efficacité des bandes dessinées sera peut-être moins mystérieuse si on l'envisage à la lumière de quelques idées au sujet des structures mentales à la base de la perception, présentées dans le premier chapitre. Nous y avons dit que le système conceptuel humain code toute donnée d'entrée en l'agençant dans une combinaison d'idéaux ou de schèmes. Ce que sous-entend cette phrase, c'est que le cerveau interprète des scènes d'après des représentations internes "idéalisées", prototypiques, qui ne correspondent jamais tout à fait à la réalité et qui doivent être rajustées au moyen de transformations pour correspondre à une scène donnée. Une théorie explique pourquoi les bandes dessineés et les caricatures sont si bien comprises: d'une certaine façon, les formes extérieures reflètent la structure interne de nos idéaux ou schèmes mentaux beaucoup plus étroitement que des photographies réalistes ou des dessins qui demeurent en étroite corrélation avec l'objet réel. C'est sans doute pour cette raison que nous sommes fascinés et amusés par les bandes dessinées et les caricatures: elles nous font percevoir la véritable nature de nos schèmes internes.

Julian Hochberg a proposé une théorie dite "de l'idéal" ou "du prototype" pour expliquer la perception des bandes dessinées. Il ne part pas toutefois des caricatures de visages ou d'êtres vivants, car elles suscitent des problèmes compliqués rattachés à l'expression. Il tente plutôt de déterminer pourquoi certains dessins semblent parvenir plus que d'autres à saisir l'essentiel de l'apparence des objets matériels. Prenons par exemple les figures 2.5 (a) et 2.5 (b). Les deux dessins sont deux projections à deux dimensions, d'égale valeur, d'un cube en fil de fer. Nous percevons pourtant plus facilement la figure 2.5 (b) comme étant un cube que la figure 2.5 (a). D'une certaine façon, la figure 2.5 (b) a saisi davantage "l'essence". Selon les termes employés par Hochberg, la figure 2.5 (b) "se rapproche davantage de la forme canonique d'un cube, c'est-à-dire des caractéristiques d'après lesquelles nous codons et nous rappelons les cubes".[1]

Hochberg allie cette notion de formes canoniques mémorisées à ce qu'il sait déjà du fonctionnement de la vision périphérique pour tenter d'expliquer pourquoi les bandes dessinées peuvent être reconnues plus facilement que les photographies dans l'étude de Ryan et Schwartz.

---

[1] Le point de vue de Hochberg est une façon d'envisager ce problème autrement que par l'approche gestaltiste classique, c'est-à-dire que la notion d'un cube tridimensionnel, une forme simple et régulière, en profondeur, est une explication plus parcimonieuse des lignes de la figure 2.5 (b). Le cerveau préfère donc voir cette représentation plutôt que l'arrangement asymétrique de lignes en tant que modèle à deux dimensions. D'autre part, la structure des lignes de la figure 2.5 (a) forme déjà un modèle régulier et symétrique à deux dimensions et résiste, de ce fait, à une interprétation tridimensionnelle.

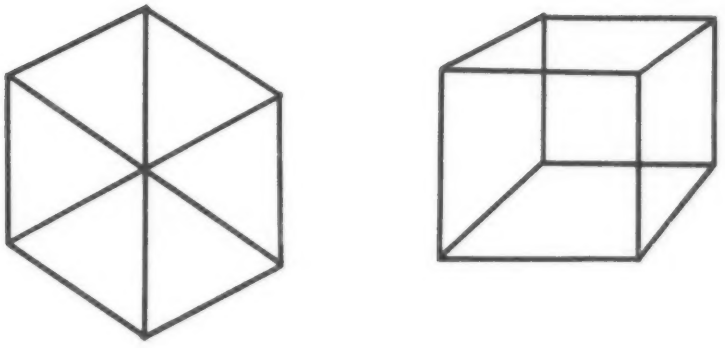


Figure 2.5. Deux projections possibles d'un cube.
Source: Hochberg (1972)

"Les contours (de la main représentée sous forme de cartoon) qui ont été retenus ont, d'une part, été "simplifiés": on a substitué des courbes régulières aux courbes complexes et irrégulières. Ce procédé fait perdre de l'information au sujet de l'anatomie de la main (l'image a donc été rendue plus redondante); on a donc besoin de moins de fixations pour l'échantillonner et faire des prédictions au sujet des parties non échantillonnées. Pour être appliquées à nos schèmes codés ou aux formes canoniques de ces objets, les caractéristiques retenues nécessitent moins de corrections (cf. Hebb, 1949; Hochberg, 1968). De plus, on peut probablement reconnaître l'objet représenté dans des zones plus excentriques du champ de la vision périphérique. C'est une autre raison pour laquelle le nombre de fixations nécessaires est réduit et celles-ci sont probablement exécutées d'après des attentes antérieures de ce qu'on verra."

Même si personne n'a effectivement vérifié dans quelle mesure ces stimuli sont reconnaissables par la vision périphérique, les raisons suivantes soustendent les affirmations ci-dessus:
1) L'augmentation de la régularité et de la redondance des courbes de la main représentée sous forme de cartoon accroît sûrement l'efficacité de la vision périphérique dont l'acuité est médiocre et qui est facilement sujette au "brouillage".
2) Chaque fois que des contours se croisent dans le cartoon, l'artiste en a délibérément modifié la disposition de sorte qu'ils se rencontrent à angle droit. Que ce soit à cause d'une "continuation satisfaisante" ou en tant qu'indice de profondeur de l'interposition, cette disposition devrait indiquer sans ambiguïté à l'observateur le bord le plus rapproché, peut-être même dans la vision périphérique.
3) Chaque fois que les contours du dessin représentent la démarcation de "trous" ou d'espaces, leur séparation relative a été amplifiée. Deux conséquences en découlent: a) chaque contour se détache plus nettement des contours voisins, même en vision périphérique, b) comme nous l'avons vu, les facteurs de proximité et de renfermement des contours tendent à transformer un espace en figure (c'est-à-dire à nous le montrer comme un objet). En augmentant la séparation entre deux doigts, par exemple, on augmente la dimension des zones qui doivent être perçues comme étant des espaces vides et, par le fait même, on empêche les espaces d'être perçues comme des objets (Hochberg, 1972, pages 75 et 76). (Traduction libre)

Hochberg estime donc que celui qui dessine un cartoon ou une caricature ne fait pas de magie, pas plus qu'il n'omet ou ne déforme de l'information déjà connue, mais qu'il effectue plutôt une sorte de traitement

préliminaire d'une image, la transformant de sorte qu'elle s'ajuste plus étroitement à la forme canonique ou idéale qu'elle a dans l'"oeil de l'esprit". C'est cette corrélation étroite qui permet à la vision périphérique de reconnaître plus facilement et plus rapidement le cartoon.

Un autre élément qui renforce une théorie axée sur la forme "canonique" ou "idéale" pour expliquer pourquoi il est plus facile de reconnaître un cartoon que son objet réel provient d'une source quelque peu surprenante: les dessins d'enfants. Arnheim (1964; 1969) a souligné que loin d'être "primitifs", les dessins d'enfants, qui font appel à un vocabulaire composé de formes simples (le cercle, la ligne droite, le rectangle, etc.) pour élaborer des représentations pictographiques, reflètent en fait la disposition naturelle du cerveau à adapter la projection rétinienne déformée (les inclinaisons, les chevauchements, etc.) aux schémas perceptifs canoniques et idéalisés. Arnheim ferait valoir par exemple que bien qu'un dessin d'enfant dans lequel un cercle est utilisé pour représenter une tête puisse ne pas paraître réaliste, le cercle rend bien cette tendance de l'esprit à coder la structure de la tête humaine comme étant essentiellement circulaire. Arnheim fournit de nombreux exemples qui démontrent comment les enfants emploient des formes abstraites simples pour expliciter et représenter, au moyen du dessin, les régularités et les simplicités sous-jacentes d'une projection rétinienne déformée qui est souvent déroutante.

Certains jeunes dessinateurs deviennent, à l'âge adulte, des artistes qui continuent d'explorer des moyens qui permettraient de donner une forme extérieure aux processus internes d'organisation de l'intelligence humaine. La caricature en est un exemple fascinant. Ici encore, l'aspect intéressant de la caricature est l'infidélité du portrait, c'est-à-dire que les caricatures semblent déroger aux règles de la projection géométrique en vue de capter l'essence d'une personne

donnée. David Perkins (1975) a présenté une analyse de la nature de la caricature qui aide à expliquer la facilité et la rapidité avec lesquelles les sujets ont compris les portraits-cartoons par rapport aux portraits-photos dans l'expérience de Ryan et Schwartz. Perkins a reproduit au trait, par calque, des caricatures de l'ex-président Richard Nixon (voir la figure 2.6). Le calquage produit de légères déviations par rapport à l'original. Les caricatures calquées demeurent pourtant plus faciles à reconnaître et résistent mieux aux effets de déviations que les calques de photographies réelles. Pourquoi? C'est peut-être que les caricatures réussissent mieux que les photographies à conserver l'ensemble de caractères distinctifs qui nous permettent de reconnaître le vrai visage de Nixon (les bajoues, le nez allongé, la coiffure, le menton carré) et résiste mieux, par conséquent, aux déformations fortuites. Ceci semble être une nouvelle confirmation du fait qu'en s'écartant du réalisme, on peut faciliter plutôt que compliquer la reconnaissance, car c'est justement en nous écartant de la réalité que nous aurions d'abord codé l'objet réel.

La théorie de l'idéal ou du schéma que nous élaborons ici concernant la reconnaissance des bandes dessinées présente cependant un problème, une sorte de paradoxe. Si nous reconnaissons les bandes dessinées et les caricatures plus facilement que les photographies et les dessins réalistes parce qu'elles reflètent l'organisation de notre schéma intérieur, comment se fait-il que nous ne prenions pas un cartoon ou une caricature pour un dessin ou un portrait réaliste? Perkins nous fournit un élément de solution à ce paradoxe dans la distinction qu'il fait entre les processus de la <u>reconnaissance</u> et de la <u>ressemblance</u>.

Figure 2.6. Calques de photographies (a) et caricatures (b) de Richard Nixon.
Source: Perkins (1975)

> La ressemblance est affaire de nuances quand on connaît l'identité du visage dépeint, soit qu'on nous la révèle, soit qu'on la reconnaisse. La reconnaissance fait appel à un système de recherche réflexive rapide qui permet d'identifier les visages. Adapté à la fonction, le système utilise relativement peu de catégorisations qui, réunies, différencient une personne d'une autre rapidement et efficacement sans exiger un examen extrêmement précis du stimulus. Après avoir reconnu le visage donné, l'observateur peut découvrir qu'il possède beaucoup d'autres renseignements à son sujet, même des renseignements contraires au stimulus.

L'exposé ci-dessus sous-entend non pas seulement la différence existant entre la reconnaissance (recherche réflexive) et la ressemblance (degré de conformité), mais aussi que l'intelligence connaît son fonctionnement perceptuel propre : le système conceptuel sait sûrement qu'il utilise des exagérations, des simplifications et des généralisations pour coder ce qu'il sait des apparences, et il peut reconnaître ces déviations pour ce qu'elles sont lorsqu'on lui présente une caricature. C'est peut-être à ce niveau de méta-sensibilisation au fonctionnement cognitif que se trouve l'explication de la raison pour laquelle nous trouvons les caricatures amusantes; les caricatures et les bandes dessinées nous rappellent, peut-être à notre insu, qu'un dessinateur perspicace a su exploiter la tendance de l'esprit humain à exagérer, à généraliser et à simplifier et a rendu ces tendances explicites. Gopnik a proposé en 1979, dans un exposé inédit, une théorie du même genre portant sur la base cognitive de l'humour. Selon Gopnik, si nous trouvons les farces drôles, ce n'est pas tant parce qu'elles se terminent de façon surprenante ou incongrue, mais plutôt parce qu'elles nous sensibilisent à la souplesse de nos schèmes intérieurs de classification qui acceptent des agencements aussi imprévus ou incongrus.

Quelles sont donc les généralisations qui se dégagent, au sujet de l'efficacité relative des photographies, des dessins, des bandes

dessinées et des caricatures? En fait, les résultats ne sont pas concluants. Les travaux de Kennedy et Dwyer confirment l'opinion voulant que les images-contours codent l'information relative aux objets aussi bien et, dans certains cas, mieux que les photographies réalistes. D'autres part, les vues de Perkins (1975) et d'Arnheim ainsi que l'expérience de Ryan et Schwartz, outre le fait qu'elles ne permettent pas de conclure à la supériorité du réalisme en tant que mode informatif, contiennent les éléments d'une théorie de la reconnaissance d'après un "idéal" ou un schéma dans laquelle les bandes dessinées ne sont pas seulement des simplifications, mais reflètent les tendances organisationnelles de l'intelligence humaine. Ceci explique qu'on les saisisse plus rapidement que les dessins ou les images réalistes. Hochberg a présenté des arguments qui corroborent l'opinion selon laquelle les bandes dessinées et les dessins-contours seraient plus informatifs que des photographies très réalistes.

Il serait commode de terminer ici en disant que, sans avoir établi clairement si les bandes dessinées sont plus informatives que les dessins, nous pouvons au moins affirmer que les images réalistes ne sont pas le meilleur moyen de présenter des renseignements. Une expérience récente de Perkins et de Hagen (1980) nous interdit cependant de généraliser de façon aussi catégorique. Au cours de leur étude, les deux chercheurs ont photographié 54 visages, tant de profil que de trois-quarts. Un artiste professionnel a ensuite fait une caricature des deux vues de chacun des 54 visages (voir la figure 2.7). Ensuite, 15 de ces 54 visages ont servi au conditionnement des sujets de l'expérience. On a montré à un groupe de sujets les caricatures de 15 visages et à un autre groupe, les photographies de ces mêmes visages. L'expérience portait principalement sur la question suivante: le fait de voir une caricature pendant leur conditionnement préparerait-il mieux les sujets à repérer les 15 visages parmi l'ensemble original de 54 visages lorsqu'on le leur présenterait au cours d'un test ultérieur? Lors du test, les visages pouvaient être présentés en photo ou en caricature, de profil ou de trois-quarts. L'expérience a démontré que

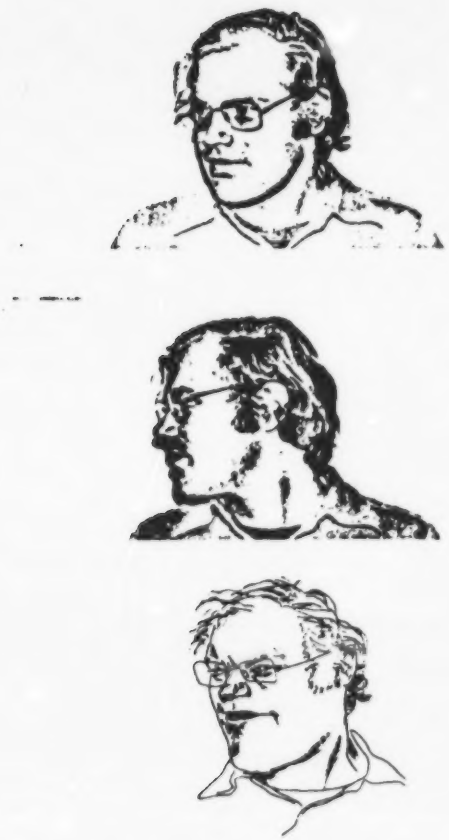


Figure 2.7. Exemples des photographies et des caricatures employées dans l'expérience de Perkins et de Hagen (1980)

le fait de voir des caricatures permet effectivement plus que la reconnaissance fortuite, et cela, même quand le jeu d'essai se compose de photographies. Cependant, les sujets à qui on avait montré les 15 photographies pendant leur conditionnement ont réussi encore mieux que l'autre groupe à les repérer parmi les 54 photographies du jeu d'essai. On voit que ces résultats compliquent le problème: non seulement ils nous mettent en garde contre la tentation de généraliser sans vérifier plus à fond la capacité informationnelle relative des photographies par rapport à celle des caricatures, mais ils remettent en question la théorie du "schéma" expliquant la mémorisation de nos concepts (du moins dans le cas des visages). Les résultats de cette expérience méritent toutefois que l'on entreprenne des études plus approfondies sur le sujet. Il y a peut-être d'autres explications du rendement inférieur du groupe auquel on a présenté des caricatures. On remarquera, par exemple, que même si elles ont été exécutées par un artiste professionnel et semblent effectivement rendre les traits distinctifs des visages, ces caricatures sont d'un style très différent des modèles de bande dessinée utilisés pour le dessin de la main dans l'étude de Ryan et Schwartz (figure 2.4). Dans ce dernier cas, les contours de la main sont plus appuyés et plus simples, comparés aux caricatures des visages représentés à la figure 2.7, où le dessin est plus subtil et plus détaillé, rappelant plutôt les caractéristiques des dessins ombrés utilisés dans l'expérience de Ryan et Schwartz. La question demeure donc entière. Aurait-il été plus facile de reconnaître les visages si les caricatures avaient été exécutées avec moins de raffinement, dans un style plus voisin de celui des dessins de Nixon (voir la figure 2.6), où les contours sont plus simples et les traits du visage, plus accentués? Jusqu'à ce qu'on entreprenne l'étude systématique de questions de ce genre, nous devons reporter à plus tard notre réponse concernant la capacité informationnelle relative des photographies et des caricatures.

2.4 Résumé et conclusions

La fonction que l'imagerie visuelle est appelée à remplir le plus souvent sur Télidon, tout au moins au début, est sans doute celle qu'Arnheim appelle sa fonction de représentation picturale (voir le chapitre I) c'est-à-dire l'utilisation d'une image pour renseigner les gens sur l'apparence matérielle des objets et des événements, par opposition à la fonction symbolique de l'imagerie, par exemple, qui vise à donner une forme précise à des idées abstraites (voir le chapitre III). Dans le présent chapitre, nous avons étudié cette fonction de représentation des images visuelles et nous en avons exploré certains aspects pouvant intéresser Télidon. Voici les principaux thèmes développés dans ce chapitre:

(1) On n'apprend pas à lire les images comme on apprend à lire les mots. Les images ne sont pas des signes arbitraires: les renseignements qu'elles fournissent sur l'apparence extérieure des objets et des événements sont présentés à peu près de la même façon que dans la réalité. Il y a cependant entre les images et les scènes réelles des différences intéressantes qui tendent à démontrer qu'on ne peut expliquer comment les gens s'adaptent au médium iconique sans tenir compte du rôle que jouent la culture et les conventions dans le processus. En général, la "lisibilité picturale" ne devrait toutefois pas créer de problème sur Télidon. Quiconque peut voir devrait être capable de voir une image visualisée sur Télidon.[1] Le

[1] Il serait peut-être plus exact de dire que quiconque peut "sentir" peut sentir une image. Cet énoncé s'inspire des travaux de Kennedy portant sur les "livres d'images pour aveugles". Ces travaux démontrent que ceux qui souffrent de cécité congénitale peuvent "voir" des dessins linéaires en relief représentant des objets familiers et des relations spatiales. Ces travaux corroborent en outre l'opinion voulant que les invariants de la structure optique soient amodaux ou abstraits par rapport à la modalité sensorielle et qu'ils puissent être rendus par des dessins linéaires.

fait de pouvoir lire une image ne garantit cependant pas que nous lirons les images exactement comme l'a prévu le concepteur, ni que deux personnes interprèteront une image de façon identique. Au-delà d'un niveau primaire d'identification (par exemple, cette image représente un escargot), une image peut faire surgir un grand nombre de propositions différentes (fragile, limoneux, lent, à antennes, etc.) selon le contexte.

(2) L'aisance et la rapidité avec lesquelles on lit une image ne sont pas nécessairement fonction du degré de réalisme photographique de l'image. Nous disposons de données, tant empiriques que théoriques, qui nous autorisent à croire possible qu'un cerveau humain, du fait qu'il interprète les données visuelles qu'il perçoit, en les ajustant à des prototypes ou schèmes mentaux, puisse comprendre plus facilement les dessins qui s'écartent de la ressemblance mécanique, comme les dessins linéaires, les bandes dessinées et les caricatures que les images photographiques. Nous avons aussi retenu l'hypothèse selon laquelle, si l'on comprend si bien les bandes dessinées et les caricatures, c'est que, dans un certain sens, ils reproduisent la forme canonique de nos schèmes mentaux. Une théorie de l'interprétation visuelle axée sur les schèmes nous a permis de formuler plusieurs hypothèses concernant les divers moyens d'exploiter les points forts de Télidon en tant que médium imagier, c'est-à-dire sa capacité de produire des dessins linéaires et des graphiques clairs et simples. Plus précisément,

(i) Est-il possible de démontrer que des dessins-contours d'objets familiers sont l'équivalent de photographies à haute fidélité dans la mesure où ils sont capables de représenter sur Télidon la plupart des variables importantes de la structure optique (forme, surfaces, arêtes, texture, fissures, etc.)?
(ii) Malgré une équivalence possible des photographies et des dessins-contours sur le plan de leur capacité de représentation

de la structure optique, les utilisateurs préfèrent-ils quand même les images de qualité photographique aux dessins linéaires concrets (si tant est que les gens croient que les photographies ne mentent pas)?

(iii) Comment la variable du délai de transmission nécessaire à la construction d'une image pictographique sur Télidon influe-t-elle sur l'acceptabilité du système? Plus concrètement, l'effet pondérateur de variables telles que le degré de réalisme photographique, le délai de transmission et le degré d'objectivité que les utilisateurs attribuent à l'image est-il tel que ceux-ci accepteraient de sacrifier les images de qualité photographique aux images-contours en retour de délais de transmission beaucoup plus courts? Quel est l'effet pondérateur de certaines caractéristiques propres à l'extraction (selon qu'on pense à la recherche en survol ou à la recherche spécialisée, par exemple)?

Ces questions concernant les effets pondérateurs des délais, du réalisme et du genre de recherche documentaire nous semblent pouvoir se prêter à une étude empirique et illustrent bien comment la recherche fondamentale portant sur les processus cognitifs et perceptuels pourrait fort bien être appliquée à la planification et à l'évaluation de nouvelles techniques de télécommunication visuelle.

que par des groupes restreints d'utilisateurs spécialisés dont les besoins relèvent de la science, de la technique, de l'éducation, de l'architecture ou d'une autre discipline.

3.0 Comment les cartes nous renseignent

Commençons par examiner une des applications les plus courantes des possibilités graphiques de Télidon: les cartes. Les cartes retiennent notre attention du fait qu'elles semblent se situer à mi-chemin du langage iconique et du langage symbolique. Il existe un certain degré d'isomorphisme entre une carte et le territoire qu'elle représente. C'est le cas, par exemple, de tout plan de ville et d'autres cartes du genre qui fournissent des renseignements sur la configuration réelle d'un territoire donné: sa forme, l'emplacement géographique des villes, certains points de repère et ainsi de suite. Cependant, une carte contient aussi certains éléments conventionnels: une ligne noire représentera une grande route, un point remplacera une ville entière, une croix désignera une église et quoi encore (voir Bertin, 1973, pour une étude détaillée du "langage" de différents types de cartes).

L'utilité des cartes provient précisément du fait qu'elles ne visent pas à représenter, dans une correspondance objective, chacun des éléments du territoire cartographié, mais qu'elles sont plutôt sélectives, ne représentant que certaines caractéristiques de ce territoire. Grâce à cette sélectivité, l'utilisateur peut effectuer facilement certains calculs à partir des renseignements qu'il tire de la carte. Il peut, par exemple, déterminer le chemin le plus court entre deux villes. Il est intéressant de voir qu'on peut tirer, sur l'efficacité des cartes, des conclusions qui sont le reflet de celles du chapitre précédent sur l'efficacité des icônes. En d'autres mots, les cartes les plus efficaces peuvent être celles qui déforment le réalisme objectif pour faciliter le processus de calcul. Arnheim (1964) nous montre un

CHAPITRE III

L'IMAGE AU SERVICE DE LA PENSÉE: DES SCHÈMES GRAPHIQUES POUR VÉHICULER DES IDÉES COMPLEXES

Des images graphiques simples et des dessins linéaires peuvent servir non seulement comme "icônes" au sens où l'entend Arnheim (voir chapitre I), c'est-à-dire comme contours dont on peut tirer des renseignements sur l'apparence physique, mais aussi comme "symboles", pour représenter les concepts ou les idées à un niveau d'abstraction supérieur à celui de l'image elle-même. Des graphiques simples (un diagramme de Venn, par exemple) peuvent nous aider à comprendre des idées complexes, spécialement dans des contextes scientifiques et pédagogiques. C'est alors qu'un schéma graphique, particulièrement s'il exploite les possibilités de transformation de l'image par ordinateur, peut supplanter le langage purement verbal sur le plan de la compréhension. (Nous ne voulons pas dire qu'un schéma graphique peut à lui seul remplacer les mots. La langue à elle seule n'est peut-être pas, non plus, le meilleur moyen de rendre un concept scientifique ou technique, mais il faudra toujours avoir recours aux mots pour orienter l'utilisateur d'un schéma graphique abstrait.) Dans cette dernière partie, nous examinons brièvement, à l'aide de quelques exemples, comment des éléments graphiques simples telles les lignes et les figures géométriques peuvent servir à des activités comme l'acquisition des connaissances et la pensée tout court. Ce qu'on doit retenir, c'est qu'il faut souvent des connaissances spécialisées pour comprendre les idées véhiculées par des éléments graphiques, même lorsque ceux-ci sont très simples. Autrement dit, contrairement à ce qui se produit dans le cas des dessins-contours ou des bandes dessinées, où l'utilisateur ordinaire de Télidon n'aura pas de peine à reconnaître les personnes ou les objets représentés, certaines images graphiques ne seront comprises

plan de métro qui en est un bon example. Le plan en format poche (voir figure 3.0) supprime les détails géographiques pour accentuer les relations topologiques pertinentes entre les points: le trajet réel des lignes du métro y est remplacé par des droites qui se coupent à 90° ou 45°, ce qui permet à l'oeil de saisir plus facilement la succession des stations et les points de correspondance. Des plans de métro schématisés comme celui de la figure 3.0 aident le cerveau humain à organiser les renseignements à mesure qu'ils lui parviennent.

Les intuitions du concepteur du plan du métro concernant la façon de déformer la réalité pour atteindre la clarté visuelle ont récemment été confirmées par des expériences psychologiques où l'on a demandé à des sujets de reproduire un dessin ou une carte d'un agencement spatial qui leur était familier, comme le plan d'un appartement (voir Norman et Rumelhart, 1975; Arnheim, 1975). Ces expériences montrent que la mémoire humaine n'est pas outillée pour reproduire des agencements spatiaux avec justesse, même lorsque ceux-ci nous sont bien connus. On voit plutôt que, dans les cartes dessinées de mémoire, les formes et les relations entre les espaces sont souvent déformées, plus droites et symétriques qu'elles ne le sont en fait, le dessin servant ainsi à traduire une réalité non pas physique, mais fonctionnelle. On en a un exemple à la figure 3.1 (b), qui illustre comment une fillette a reproduit de mémoire l'agencement physique de l'appartement de ses parents. La disposition réelle des pièces est indiquée à la figure 3.1 (a). La reproduction faite par la fillette, comme c'était le cas du plan du métro, révèle la tendance à faire ressortir les relations topologiques et fonctionnelles entre les espaces pour rendre les renseignements visuels plus clairs. Le plan dessiné par la fillette semble représenter, mieux que le plan techniquement exact qui montre simultanément les multiples relations entre les éléments, l'expérience temporelle et linéaire de quelqu'un qui se déplace autant que possible en ligne droite dans un milieu donné.

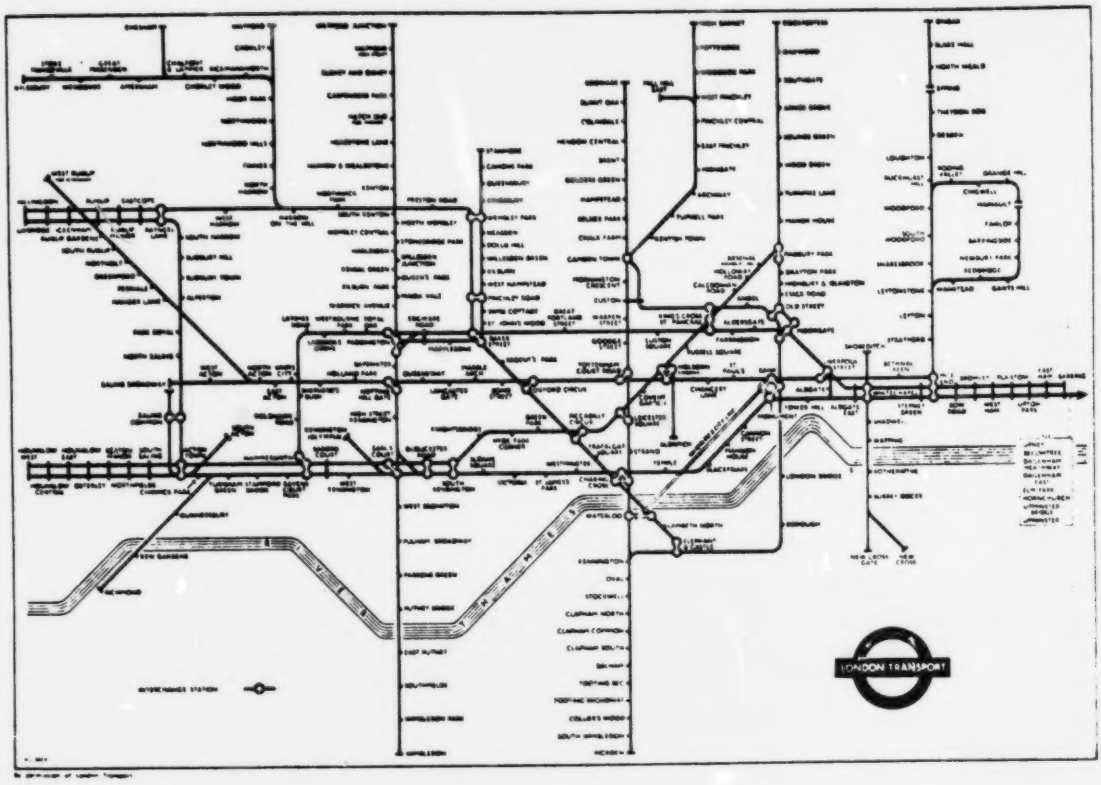


Figure 3.0 Plan de métro

Source: Arnheim (1964)

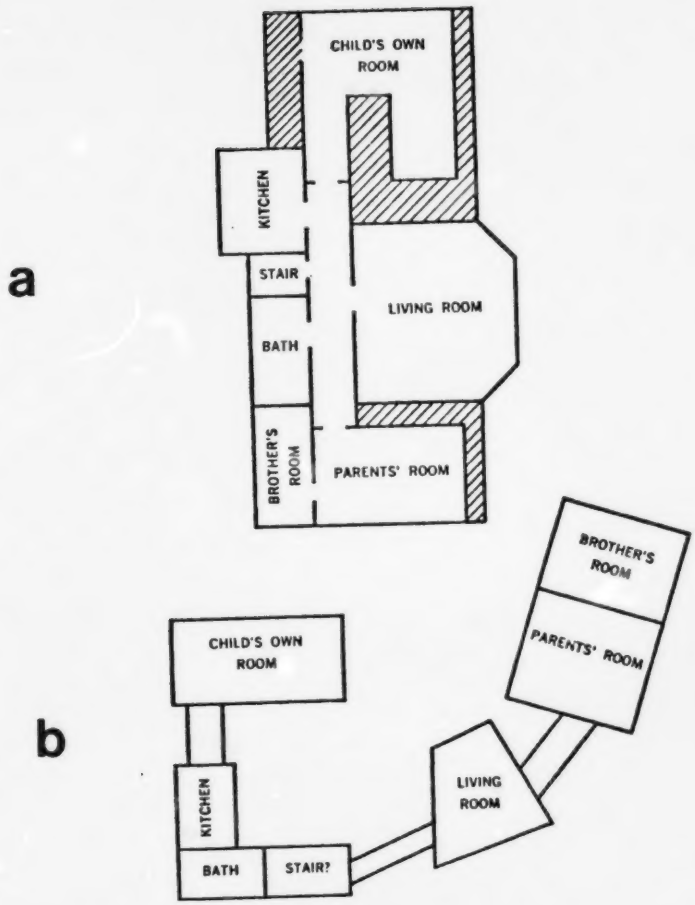


Figure 3.1 Disposition réelle des pièces d'un appartement (a) et disposition de ces pièces reproduite de mémoire par une fillette de douze ans (b).

Source: Arnheim (1975)

3.0.1 Aptitude à lire les cartes: différences entre les individus

Télidon peut facilement reproduire toutes sortes de cartes, notamment les cartes topographiques, vu leur simplicité. Mais quels mécanismes interviennent dans l'acquisition de connaissances à partir d'une carte? La lecture d'une carte est-elle aussi facile que celle d'une image? La plupart des utilisateurs sauront-ils comprendre l'information véhiculée par des cartes visualisées sur Télidon? L'aptitude à lire, mémoriser et utiliser l'information contenue sur une carte ne varie-t-elle pas plutôt selon les individus? S'il existe effectivement des différences, sont-elles dues à l'expérience et peuvent-elles dans ce cas être supprimées par la formation ou sont-elles reliées à des facultés cognitives innées? Une étude récente de Thorndyke et Stasz (1980), bien qu'elle ne porte pas sur les cartes visualisées sur Télidon, tente de répondre à ces questions.

Dans la première partie de leur expérience, les auteurs ont demandé à des "experts" et à des "profanes" de mémoriser le plan d'une ville fictive de façon à pouvoir le reproduire de mémoire et résoudre des problèmes de navigation, comme déterminer le meilleur itinéraire entre deux points, par exemple. Les participants s'exprimaient à haute voix pendant les épreuves que comportait leur "formation". Un système de notation fondé sur l'exactitude des reproductions a permis aux auteurs de distinguer les "bons élèves" des "élèves médiocres". Ils ont constaté que les premiers suivaient des stratégies ou des "méthodes efficaces" pour assimiler les relations spatiales, évaluer leurs progrès et se concentrer sur les renseignements non assimilés. De façon inattendue, cette expérience a montré qu'il n'existait pas de relation entre la facilité d'assimilation et le fait de posséder de l'expérience en la matière. Ce qui importait, c'était le recours à des "méthodes efficaces".

Dans une expérience ultérieure, les auteurs ont voulu savoir si l'enseignement de "méthodes efficaces" à des profanes reconnus aussi capables les uns que les autres d'apprendre à assimiler des cartes allait accroître leur rendement par rapport à celui d'un groupe témoin auquel on enseignait des "méthodes neutres". Les auteurs ont de fait constaté qu'il est possible d'enseigner à des individus à mieux assimiler une carte en leur enseignant des méthodes efficaces. Cependant, l'étude a donné lieu à une autre constatation intéressante: une fois l'expérience achevée, on a soumis chacun des sujets à un test de "mémoire visuelle" pour déterminer l'aptitude à se rappeler correctement des éléments spatiaux. On a alors découvert qu'il y avait interaction entre la formation et la mémoire visuelle. Plus précisément, les sujets qui avaient une "bonne mémoire visuelle" apprenaient plus facilement les "méthodes efficaces" que les sujets qui avaient une "mauvaise mémoire visuelle". En fait, ceux-ci ne s'amélioraient pas plus que les sujets du groupe témoin auquel des "méthodes efficaces" n'étaient pas enseignées.

Cette expérience permet de dégager certaines questions pratiques concernant l'utilisation de cartes sur Télidon. Tout d'abord, le fait de présenter une carte - même un plan de ville ordinaire - ne signifie pas que n'importe qui pourra facilement l'interpréter (spécialement s'il faut mémoriser des éléments de la carte pour s'en servir ultérieurement). Même si les cartes sont constituées à partir d'éléments graphiques simples, l'interprétation des renseignements qu'elles contiennent peut faire appel à diverses facultés mentales que tous ne possèdent pas dans les mêmes proportions. Il semble que le processus d'assimilation des cartes soit nettement individualisé. Cependant, la cause de ces différences d'un individu à l'autre est la question qui nous préoccupe davantage. Ces différences sont-elles dues à un manque d'expérience ou résident-elles dans les facultés cognitives innées? S'il faut avoir une bonne mémoire visuelle pour apprendre à

bien se servir d'une carte, l'enseignement de méthodes efficaces pour aider les individus à tirer parti d'une carte pourra être une perte de temps, sauf si l'on trouve une technique d'enseignement pour améliorer la mémoire visuelle (qui semble une aptitude innée). L'utilité des cartes serait alors moins générale qu'on aurait pu le croire. C'est peut-être, en partie, qu'une carte, comme les bleus ou les plans d'un architecte, montre les objets et la façon dont ils sont disposés en plongée ou selon d'autres points de vue, comme on les voit rarement dans le cours normal de l'expérience perceptuelle quotidienne. Il est également important de souligner que la lecture d'une carte ou d'un bleu, contrairement à celle d'une image, ne vise pas uniquement à en reconnaître le signifié, mais aussi à engager le processus de traitement de l'information spatiale pour résoudre un problème quelconque, par exemple naviguer entre des points, estimer des distances ou identifier les traits caractéristiques d'un terrain. Or la capacité de résoudre de tels problèmes peut dépendre non seulement de la clarté visuelle de la carte, mais encore de la possession de certaines aptitudes à interpréter l'information apparemment enracinées dans des facultés cognitives fixes. Il faudrait en connaître plus long sur les blocs de construction cognitifs qui servent à enseigner à se servir d'une carte. Des recherches dans ce domaine devraient jeter plus de lumière sur l'utilisation rationnelle des cartes dans des systèmes de recherche documentaire comme Télidon.

3.0.2 Cartes cinématographiques

L'expérience de Thorndyke et Stasz indique qu'il peut exister, d'une personne à l'autre, des différences d'aptitudes, enracinées dans les facultés cognitives fixes, à tirer des connaissances des cartes et à les utiliser. Mais une carte "topographique" ou "vue en plongée" est-elle le seul instrument dont dispose l'utilisateur pour acquérir des connaissances sur un environnement spatial? Quelqu'un pourrait peut-être, à l'aide des nouvelles techniques de graphisme nées de l'informatique, concevoir d'autres types de cartes pour tirer profit de

la connaissance qu'ont les individus de leur univers perceptuel quotidien. Pour suppléer aux plans de ville habituels, l'Architecture Machine Group du MIT examine actuellement une solution novatrice qui réunit les possibilités du graphisme informatisé et celles du vidéodisque optique. Le projet de recherche a été intitulé "carte cinématographique" (Lippman, sans date) et voici comment Lippman en décrit l'idée sous-jacente:

> ... Sous sa forme la plus élémentaire, on peut la considérer (carte cinématographique) comme un substitut dynamique de la carte topographique en papier: elle permet à l'utilisateur, ou à celui qui la "lit", de se familiariser avec un environnement spatial. Cependant, cette familiarisation s'accomplit par des moyens fort différents: le "lecteur" explore l'espace cartographié en participant à une promenade simulée au cours de laquelle il voit des séquences filmées qui reproduisent ce qu'il verrait s'il parcourait cet espace en voiture. La sensation de la conduite automobile est intensifiée et le "conducteur" y est d'autant plus associé qu'il y a interaction: c'est lui qui détermine les parcours, les virages, la vitesse et les points d'observation. Il peut aussi choisir la saison au moyen d'un bouton "saisons" et le mode visuel de l'excursion: photo, croquis ou dessin animé. Ainsi, la carte cinématographique n'est pas la simple réplique d'une visite guidée, mais permet au "visiteur" de partir librement à la découverte, d'aller à son propre rythme et d'emprunter les voies qu'il désire. (Lippman, page 4)

Outre l'excursion simulée dans un environnement spatial, la carte cinématographique permet à l'utilisateur d'obtenir des renseignements utiles sur les immeubles, les rues et autres points d'intérêt qu'il verrait en cours de route. Par exemple, un conducteur pourrait s'arrêter devant un restaurant chinois et demander la carte. On pourrait donc concevoir une carte cinématographique qui serait aussi une sorte d'index iconique dynamique.

La carte cinématographique fait appel au stockage, sur une paire de vidéodisques, d'images préalablement filmées (sous plusieurs angles en même temps) lors d'une promenade en ville. Chaque vidéodisque peut stocker 54 000 images de télévision auxquelles l'utilisateur peut avoir directement accès. Comme l'adresse de chaque image peut être codée sous une forme que peut déchiffrer l'ordinateur, il est possible de créer des programmes qui laissent à l'utilisateur beaucoup de liberté dans le "montage" des images codées: constitution de séquences d'images, délimitation des vitesses de déplacement, etc. L'utilisation de deux vidéodisques permet au "conducteur" de choisir au fur et à mesure qu'il progresse. Il peut, par exemple, décider de tourner à gauche à l'intersection suivante. En effet, pendant qu'un vidéodisque présente les images simulant le déplacement dans une rue, l'autre, le vidéodisque "d'anticipation", est automatiquement réglé pour que le défilement de ses propres images de la rue commence à l'intersection. Lorsque l'utilisateur indique un virage, le défilement des images du deuxième vidéodisque s'amorce. Une fois le virage effectué, le premier vidéodisque, dont les images ont entretemps été remises en position, entre en jeu. L'excursion peut donc se poursuivre de façon ininterrompue à l'intersection suivante.

La carte cinématographique possède de nombreuses autres caractéristiques novatrices (signaux de contrôle des séquences commandées par le toucher et synthèse du son, par exemple), mais il serait trop long de les passer toutes en revue. Cependant, avant de passer à un autre sujet, nous aimerions contester une des hypothèses fondamentales de ses créateurs. Nous avons vu que l'objet d'une telle carte était de remplacer la carte topographique en papier d'une région par une promenade dynamique dans cette région. La solution était intéressante compte tenu du problème que peut présenter la disparité des aptitudes de chacun à lire les cartes, suivant l'étude de Thorndyke et Stasz.

Autrement dit, certaines personnes peuvent avoir de la difficulté à interpréter une carte topographique, mais tout le monde devrait pouvoir se servir d'une carte cinématographique puisqu'elle reproduit, jusqu'à un certain point, ce qu'un utilisateur verrait s'il se trouvait sur place. Cependant, une carte cinématographique peut-elle vraiment remplacer un plan de ville? Est-elle vraiment conçue pour remplir la même fonction? Même si un plan de ville, format poche, ne rend pas l'aspect des maisons, des immeubles et des points de repère, il fournit à l'utilisateur une vue globale du territoire, c'est-à-dire les relations spatiales entre les rues et les points de repère. Muni d'un tel plan, il peut trouver où il est, où il veut aller et le meilleur chemin pour s'y rendre. Le réseau de connexions y est présenté en une seule image de l'espace qu'il recouvre. Un plan de ville ne vous indique pas, cependant, à quoi vous reconnaîtrez le musée situé au coin de la rue du Parc et de la rue Principale une fois parvenu à cette intersection. Comment pourriez-vous, si vous n'aviez que cette carte cinématographique, trouver le chemin qui vous mènera au musée si vous n'y êtes jamais allé? La seule façon serait d'effectuer le trajet. De plus, si cela vous oblige à faire des tours et des détours, un seul parcours ne suffira peut-être pas, car vous devrez mémoriser une foule d'impressions séquentielles et temporelles. On peut donc dire qu'une carte cinématographique est linéaire et séquentielle dans sa façon de faire percevoir des relations spatiales entre les éléments alors qu'une carte topographique, au contraire, donne à voir simultanément les multiples connexions du réseau spatial. Pour cette raison, cette carte cinématographique ne pourrait pas se substituer avantageusement à une carte topographique. Cependant, elle pourrait venir la compléter. De fait, les concepteurs de la carte cinématographique ont déjà prévu cette possibilité en travaillant à la création d'effets d'animation informatisés pour "permettre une vue d'ensemble interactive de la ville

et supprimer les détails superflus qui risquent d'embrouiller l'utilisateur". L'idéal serait de pouvoir faire des "excursions" simulant la réalité entrecoupées à volonté de vues d'ensemble de la ville comme en fournissent les cartes. Même si l'accès à pareil système ne pourra probablement pas être généralisé dans un proche avenir, la nature de la carte cinématographique montre pourquoi les communications imagières constitueront sûrement, un jour, une partie importante des interactions réalisées grâce aux ordinateurs et - qui sait? - par une version future de Télidon.

3.1 Schèmes graphiques et pensée

Ceux qui travaillent dans les universités et dans les milieux scientifiques ont souvent trouvé, en entrant dans un amphithéâtre ou une salle de conférences, un tableau noir ou un tableau à feuilles détachables rempli de griffonnages, cases, flèches, cercles ou autres diagrammes, parfois accompagnés d'inscriptions. Un jeu amusant consiste à deviner, à partir des seules représentations graphiques, le sujet discuté. La plupart du temps, ces dessins fournissent des indices sur les propriétés constitutives du sujet, quoique le sujet lui-même ne soit pas facilement identifiable. Nous reconnaissons ordinairement dans les gribouillages les relations "pures" qu'ils étaient destinés à signifier même si nous ne connaissons pas exactement la nature des éléments en cause. Pensons, par exemple, aux concepts de "chevauchement" (diagramme de Venn), de "cheminement" ou de "transmission" (organigramme), de "hiérarchie" (dendrogramme) et ainsi de suite. Ces schémas traduisent des caractéristiques structurales qui peuvent s'appliquer à différents domaines: une "flèche" peut se retrouver dans un programme informatique, un schéma de circuit ou l'organigramme d'une entreprise. Pourquoi les enseignants et les scientifiques éprouvent-ils le besoin de griffonner

ces graphiques? Si la pensée est, comme certains le croient, un "langage interne", pourquoi ne pas s'exprimer à l'aide de mots uniquement? Ces cases et ces flèches rudimentaires viennent-elles réellement en aide à la pensée? Ne serviraient-elles pas plutôt à attirer l'attention ou ne seraient-elles que des conventions, sans lien réel avec les pensées abstraites qu'elles sont censées illustrer?

Les psychologues ont toujours cru que la pensée prend soit la forme d'un "langage interne", soit celle "d'images mentales", ces dernières n'étant que de pâles reproductions imagières de l'expérience perceptuelle (Paivio, 1971). Bien que notre étude n'ait pas pour objet la rédaction d'un traité sur "le langage de la pensée" (on peut quand même consulter Fodor qui en a traité en 1980), nous devons préciser que quelques savants ont retenu une troisième hypothèse qui pourrait expliquer pourquoi les "schémas graphiques abstraits" soutiennent la pensée. Rudolph Arnheim, par exemple, s'est fait l'ardent défenseur de la notion selon laquelle <u>toute</u> pensée est "perceptuelle", terme qu'il oppose à "iconique". Selon Arnheim (1969), ce ne sont ni les mots, ni les images mentales iconiques qui sont le support de la pensée, mais une certaine imagerie abstraite qui rend sous une forme commode pour la pensée les propriétés "génériques" utiles des objets, des situations et des événements. Pour Arnheim, le rôle de la langue dans la pensée est nettement subsidiaire à celui de la connaissance perceptuelle. Il estime qu'on ne peut <u>pas</u> penser en mots. Ceux-ci ne sont que les sentiers qui aboutissent au véritable support de la pensée agissante, l'imagerie abstraite des formes, des relations et des forces génériques. Les mots ont pour principal rôle de mobiliser la connaissance perceptuelle abstraite qui sous-tend la langue et de déclencher les processus mentaux dont cette connaissance fait l'objet.

Il n'est pas possible d'observer directement les images abstraites de la pensée, mais nous pouvons nous faire une idée de leurs attributs en examinant les résultats d'une expérience dont Arnheim fait état dans son ouvrage Visual Thinking. Dans cette expérience, on a demandé à des gens d'imaginer des métaphores graphiques pour exprimer des concepts "abstraits" comme "le passé, le présent et le futur", la "démocratie" et celui "de mariage réussi/mariage raté". Si différents sujets ont rarement imaginé les mêmes symboles graphiques pour illustrer ces concepts, ils ont néanmoins trouvé facilement certaines images schématisées qui rendaient, selon eux, certains aspects de ces concepts. La figure 3.2 montre, par exemple, de quelle façon un sujet a illustré le concept de "mariage réussi/mariage raté". L'intérêt de ce dessin n'est pas qu'il exprime bien tout le sens que véhiculent les mots "mariage réussi/mariage raté" mais plutôt qu'en dépit du caractère abstrait et schématique de ces images (rien ici ne représente concrètement les conjoints, le gâteau de noces ou le rouleau à pâtisserie), nous pouvons quand même voir comment les caractéristiques structurales des formes expriment, sur un plan abstrait, quelques différences importantes qu'on fait entre un mariage réussi et un mariage raté. Nous pourrions paraphraser les différences exprimées par ces dessins en disant: "Dans un mariage réussi, il existe deux entités indépendantes et pourtant elles s'emboîtent pour former un tout solide; dans un mariage raté, par contre, les deux éléments distincts ne s'enclenchent pas, ils ne sont que faiblement joints l'un à l'autre, à

Figure 3.2 Métaphore visuelle du concept "mariage réussi/mariage raté"

Source: Arnheim (1969)

un niveau superficiel, pour former un tout moins résistant".[1] Soulignons encore une fois que l'essentiel n'est pas qu'un schéma graphique puisse servir à rendre des concepts complexes comme le "mariage", la "démocratie" ou autres: nous avons déjà les mots voulus pour cela. Ce qui importe, c'est que les représentations schématiques abstraites peuvent donner une forme visible à des constellations de forces et de relations (par exemple, "interdépendance") qui peuvent nous aider à mieux concevoir ces relations. Je crois que c'est en partie pour cette raison qu'ingénieurs, éducateurs et scientifiques trouvent ces schémas graphiques utiles. Tout comme la métaphore visuelle de la figure 3.2 nous aide à structurer notre pensée relativement à un concept comme le "mariage", un schéma fonctionnel ou schéma de circuit peut remplir le même rôle dans les domaines scientifiques et techniques en éliminant des détails superflus et en faisant ressortir des caractéristiques et des relations structurales importantes (inclusion, chevauchement, cheminement, etc.) dont l'occurrence n'est pas directement observable.

Le "dendrogramme" (diagramme à structure ramifiée) illustre bien comment un schéma graphique donne une forme visible à des notions complexes. La façon dont un dendrogramme donne à voir les liens hiérarchiques logiques qui existent entre divers éléments d'un ensemble dépasse la capacité qu'aurait la langue de transmettre seule la même information. Imaginons qu'il faille reconstituer les ramifications d'un arbre généalogique ou les éléments d'un organigramme d'après une description verbale (Johnson relève de Adams, qui relève de Clark, qui relève de Jenkins, qui relève de ...). Un autre exemple courant nous

---

[1]Le fait que nous soyons capables de comprendre de telles métaphores visuelles malgré l'absence de caractéristiques extérieures communes aux images et aux mots indique que la connaissance linguistique et la connaissance perceptuelle se rejoignent sur un plan plus abstrait que celui des mots ou des images réalistes concrètes. On retrouve des opinions semblables sur le rôle de l'imagerie schématique abstraite dans la langue et la pensée chez Verbrugge (1974), Lakoff (1980) et Mills (1980).

est fourni par les ordinogrammes, dont la structure visuelle nous aide à "voir" les relations entre les parties d'un programme informatique complexe avec une clarté que ne permettrait pas la lecture des instructions en langage de programmation.

3.1.2 Graphisme dynamique

Dans la section précédente, nous avons laissé entendre qu'une image graphique complexe n'était pas nécessairement le meilleur moyen de rendre une idée complexe. Nous avons dit, au contraire, que si les scientifiques, les éducateurs et les ingénieurs ont souvent recours à des graphiques schématiques, ce n'est pas à cause de leur complexité, mais de leur simplicité. Ces graphiques donnent une forme à des propriétés et à des relations abstraites dont l'occurrence n'est pas directement observable. Des diagrammes schématiques de toutes sortes, facilement réalisables grâce aux possibilités graphiques actuelles de Télidon, pourront répondre à quantité de besoins spécialisés. Nous n'avons présenté jusqu'à maintenant que des exemples de schémas comme on en trouverait sur un tableau ou dans un manuel, c'est-à-dire des images statiques, figées. Dans cette dernière partie, nous allons chercher à déterminer si un graphique ou une image fixe exploite pleinement les possibilités de l'ordinateur (et donc de Télidon) en tant qu'instrument servant à produire des graphiques pouvant venir en aide à la pensée. Autrement dit, dans quelle mesure le fait de voir les images bouger ou se transformer pourrait-il faciliter la compréhension des idées?

Il faut, pour résoudre des problèmes et acquérir une connaissance, aller au-delà de l'information qui nous est fournie. En d'autres termes, pour que la réflexion donne des résultats, il faut "restructurer" la perception initiale que l'on a d'un problème, d'un événement ou d'une situation afin de découvrir les pistes qui conduisent à une issue (Newell et Simon, 1972). Cette opération de "restructuration" est en elle-même dynamique puisqu'il nous faut changer ou modifier le modèle

interne initial d'une situation, c'est-à-dire le faire passer d'un état de connaissance à un autre état. Or, si une image fixe ou un schéma graphique peut faciliter cette opération de restructuration, des images qui subissent des transformations pourraient accomplir la même fonction encore plus efficacement. L'ouvrage d'Arnheim intitulé Visual Thinking nous fournit un exemple. Examinons la figure géométrique présentée à la figure 3.3. Le problème consiste à en trouver la superficie lorsque A et B sont connus. À prime abord, le problème peut sembler difficile parce que nous avons tendance à percevoir cet agencement géométrique au niveau structural le plus simple: un parallélogramme qui se superpose, en oblique, à un carré. Cependant, on perçoit plus facilement la solution si l'on considère que les lignes représentent un autre type de construction, c'est-à-dire deux triangles rectangles qui se chevauchent et qui ont chacun une superficie de $\frac{1}{2}$ (A-B). Même si le stimulus verbal peut suffire à déclencher le processus de "restructuration" requis, la démonstration serait beaucoup plus convaincante si le mode de présentation consistait en un film d'animation: on observe la figure qui, se séparant lentement en deux, révèle chacun des deux triangles qui se déplacent comme des entités distinctes sur l'arrière-plan. Ensuite, les deux triangles se rapprochent de nouveau, puis se chevauchent pour former l'image fixe de la figure 3.3. La formation de pareilles séquences dynamiques pour faciliter l'enseignement de concepts mathématiques ou scientifiques est possible non seulement au moyen des techniques courantes du cinéma d'animation, mais aussi en faisant appel aux possibilités de graphisme informatisé des ordinateurs modernes. De plus, l'ordinateur possède le potentiel voulu non seulement pour donner vie à des images, mais aussi pour conduire à une interaction personne-image, c'est-à-dire permettre à l'utilisateur de faire des choix et d'en percevoir les conséquences (ainsi, en 1976, Lutz et Rigney ont montré que le graphisme informatisé interactif est supérieur à la présentation langagière pour l'enseignement de concepts élémentaires d'électro-chimie).

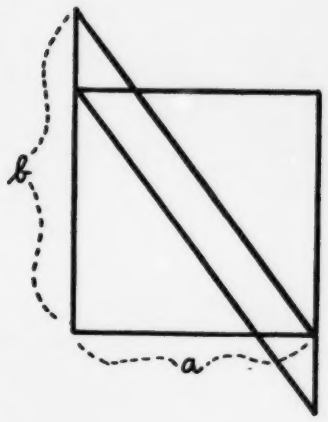


Figure 3.3 Solution d'un problème géométrique

Source: Arnheim (1969)

Le livre de William Huggins et Doris Entwistle intitulé Iconic Communications (1974) est une des tentatives les plus ambitieuses en vue d'évaluer le potentiel de communication du graphisme dynamique. Ces auteurs estiment qu'avec l'avènement de l'ordinateur en tant qu'outil de stockage et de manipulation des images, le temps est venu d'instituer une nouvelle science, l'iconique, qui canaliserait les connaissances en éducation, linguistique, communications, psychologie et informatique pour les appliquer à la production d'images dynamiques à des fins éducatives. Même si les auteurs ont tendance, selon moi, à tomber dans le piège de chercher à substituer des images aux mots (voir Chapitre I du présent rapport), ils apportent quelques exemples opportuns et font état d'expériences qui permettent de mieux voir comment le graphisme dynamique peut faciliter la compréhension de notions en éducation. Leur ouvrage contient en outre une abondante bibliographie annotée.

Un des exemples fournis par les auteurs est particulièrement pertinent: ils ont tenté d'établir une notation de "symboles dynamiques" - signes iconiques arbitraires qui changent d'état - pour enseigner certaines notions élémentaires du passage d'un courant dans un circuit électrique: pourquoi certains composants sont dits actifs (accumulateurs) tandis que d'autres sont dits passifs (résistances); comment le courant se divise lorsqu'il peut emprunter plusieurs branches; comment la disposition topologique des composants influe sur le passage du courant (composants montés en parallèle ou en série). Comme le signalent les auteurs, la difficulté réside ici dans le fait qu'on ne peut pas, simplement, filmer l'apparence d'un accumulateur et de quelques résistances. Même si un accumulateur et une résistance n'ont pas la même apparence, leurs différences matérielles n'aident pas à en comprendre les propriétés électriques. Les auteurs soulignent en outre que l'illustration de ces différences par les moyens classiques, c'est-à-dire l'utilisation d'indicateurs mécaniques ou de voltmètres numériques, par exemple, n'exploite pas les possibilités qu'offrirait l'iconique de montrer les relations structurales et causales.

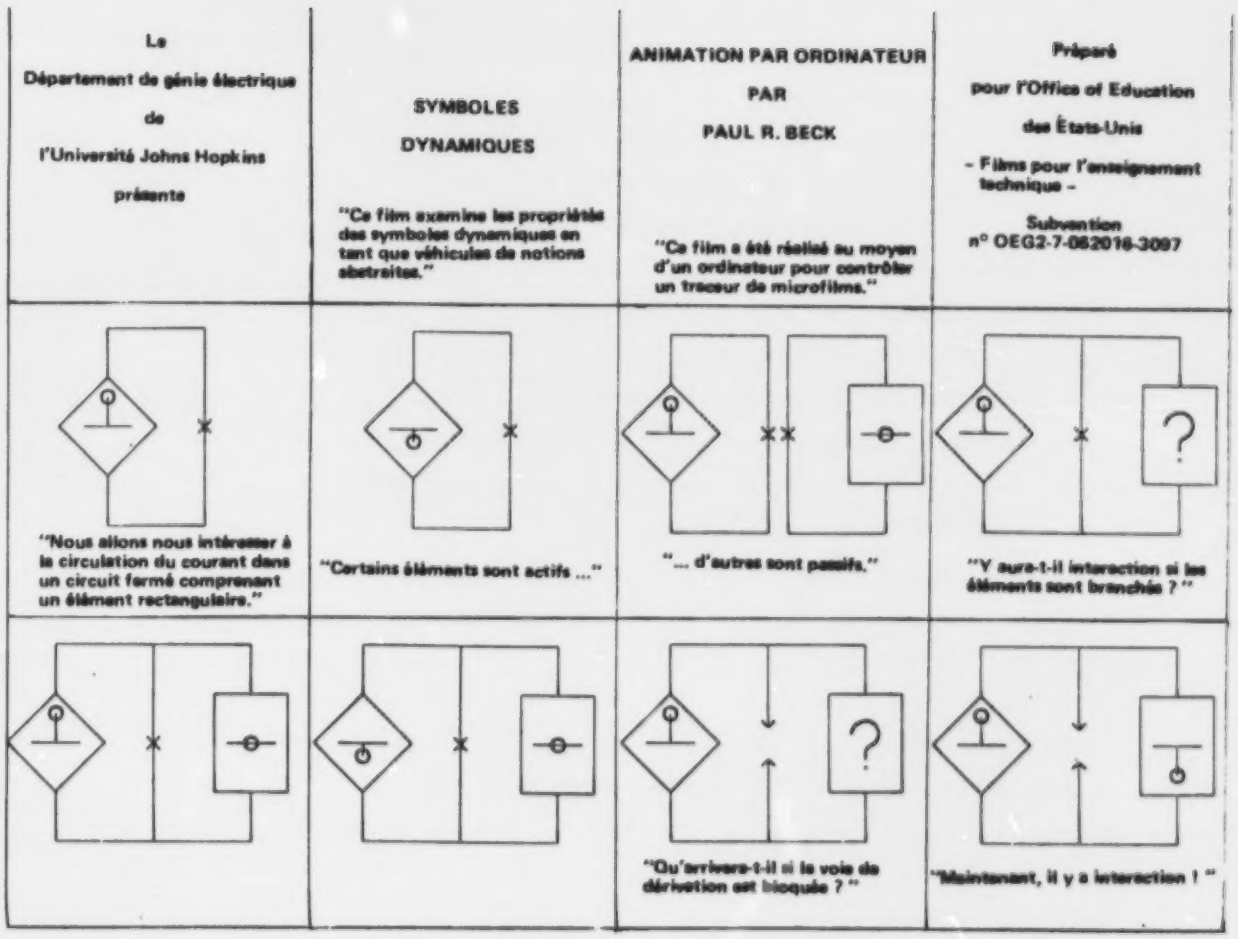


Figure 3.4 Images fixes tirées du film sur les "symboles dynamiques".

Source: Huggins et Entwistle (1974).

En guise de solution, on a réalisé un film où les composants d'un circuit électrique et leurs propriétés fonctionnelles sont représentés par des "symboles dynamiques". La figure 3.4 montre plusieurs images de ce film. Les auteurs ont délaissé la symbolique classique des manuels pour représenter les circuits par une notation plus simple qui, espèrent-ils, rendra plus naturellement les idées essentielles. Par exemple, au lieu du symbole qui représente ordinairement le voltmètre, celui qu'ils emploient ressemble à une "sucette" ronde dont on voit rallonger ou raccourcir la tige par rapport à une ligne de référence selon que l'intensité du courant augmente ou diminue. Le film vise à illustrer des notions élémentaires de débit du courant, de résistance et d'additivité dans un circuit électrique en montrant les changements apportés à ces éléments graphiques simples. Dans un test subséquent portant sur la valeur pédagogique du film, les auteurs montrèrent que tant des étudiants en génie que des étudiants dans des disciplines autres que les sciences, avaient mieux réussi qu'un groupe témoin à saisir des notions élémentaires d'électricité. Cependant, une différence intéressante existaient entre les groupes. Les étudiants en génie ont interprété le film selon la terminologie de la théorie des circuits, tandis que les étudiants dans d'autres disciplines ont "anthropomorphisé" les notions (d'une façon quelque peu sexiste). Par exemple, ils ont décrit un accumulateur comme un "homme" (composant actif) qui essaie de communiquer avec une "femme" (résistance passive). Le fait de voir des symboles dynamiques comme des métaphores visuelles représentant des actions sociales témoigne, encore une fois, de la capacité qu'ont les graphiques abstraits d'exprimer des caractéristiques structurales générales dont on peut trouver des exemples dans de nombreux domaines. Par exemple, le dessin animé permet d'observer des modèles communs aux phénomènes scientifiques et aux phénomènes sociaux. Finalement, les auteurs, ont noté que si ces simples graphiques animés

semblent constituer un succès, ces films ne sont pas faciles à monter sur le plan conceptuel, même si leur réalisation technique ne présente pas de difficultés. En effet, il faut beaucoup de temps et de talent pour traduire sous forme de graphiques dynamiques des notions conceptuelles, surtout lorsqu'on touche au domaine scientifique. Les applications futures de graphiques abstraits à des fins pédagogiques au moyen de Télidon seront soumises à ces contraintes d'imagination visuelle.

3.1.3 Mouvement et séquence

Une dernière remarque au sujet des graphiques dynamiques s'impose. Voir une image ou un graphique en voie de transformation peut aider à restructurer les opérations qui sont nécessaires à la pensée agissante. De plus, la communication de changements et de transformations peut s'accomplir par des séquences animées, c'est-à-dire des images fixes qui se succèdent à une vitesse qui donne au spectateur l'illusion d'un mouvement régulier. (Il est actuellement possible de réaliser sur Télidon des séquences animées d'un mouvement continu, quoique la création de l'apparence du déplacement au moyen du système actuel de composition des pages semble quelque peu laborieuse.) Au cinéma ou à la télévision, on crée l'illusion du mouvement continu grâce à la transmission de 24 ou 30 images par seconde. Mais le mouvement "apparent" est-il toujours essentiel pour comprendre un événement dynamique? Pas nécessairement. Je suis d'avis qu'il est opportun d'établir une distinction entre la perception du "mouvement" et la compréhension de la "séquence". Pour comprendre la signification d'une séquence dynamique, il faut reconnaître les changements et les transformations d'état qui caractérisent certaines actions (marcher, fondre, entrer en collision, se rompre, etc.). La spécification de ces

transformations n'exige pas l'illusion du mouvement continu, mais seulement la capacité d'intégrer l'information provenant de plusieurs échantillons successifs d'images fixes, dont chacun représente un état discret d'un événement. C'est ce qui se produit pour la bande illustrée ordinaire. Ainsi, il n'est pas nécessaire de percevoir un mouvement continu pour comprendre que la bande dessinée de la figure 1.1 utilise quatre images discrètes pour rendre l'idée d'un événement dynamique, c'est-à-dire un objet qui "avance" vers l'observateur. Ce qu'il faut, c'est connaître la syntaxe des bandes illustrées. De même, on pourrait facilement interpréter le film qui utilise des "symboles dynamiques" dont nous venons de parler comme une bande dessinée ou même un diaporama où des images fixes se succèdent sur l'écran de l'utilisateur.

La distinction entre l'aptitude de l'esprit à comprendre la signification d'une "séquence" et l'aptitude du système perceptuel à saisir la notion de "mouvement apparent" laisse supposer qu'on pourrait, pour présenter des données sur des événements visuels dynamiques, examiner des méthodes qui s'appuieraient sur des "échantillons" d'images fixes. Pareilles méthodes permettraient de réduire sensiblement les largeurs de bande nécessaires aux transmissions et se traduire par une économie appréciable d'espace mémoire dans le cas d'un codage numérique de séquences visuelles dynamiques (voir Mills qui, dès 1972, a traité de cette question).

3.2 Résumé et conclusions

Un des principaux thèmes qui ressort du présent document est qu'une image, pour être significative, n'a pas besoin d'être une réplique fidèle de la réalité. Dans ce chapitre, nous avons vu quelques moyens d'utiliser des schémas graphiques, c'est-à-dire des images très abstraites utilisant des éléments graphiques simples, pour comprendre

plus facilement des concepts complexes, y compris la configuration d'environnements spatiaux et l'enseignement de notions mathématiques et scientifiques. Voici un résumé des principaux points de ce chapitre.

1) L'utilisation de cartes est une des applications de la représentation graphique qu'on cite le plus souvent. On a pu constater que les cartes les plus efficaces ne sont pas nécessairement celles qui sont le plus réalistes, mais celles qui, de fait, "déforment" la réalité en éliminant des détails et en clarifiant, pour l'oeil, les liens topologiques et fonctionnels entre les entités géographiques (un plan de métro, format poche, par exemple).

2) En dépit de l'utilité évidente des cartes, l'aptitude à lire une carte et à en tirer les connaissances voulues peut varier d'un individu à l'autre. Les "bons élèves" utilisent des "méthodes efficaces" (stratégies de traitement de l'information) pour assimiler une carte et s'en servir pour résoudre des problèmes. Les "élèves médiocres" n'ont pas recours à ces méthodes. De plus, l'aptitude à apprendre des "méthodes efficaces" peut être liée à des capacités cognitives innées comme la mémoire visuelle. En ce qui concerne Télidon, ce qui importe, c'est que, contrairement aux représentations d'objets sous forme de bandes dessinées, de caricatures ou de croquis que chacun, en principe, interprète facilement, certains types de représentations picturales (cartes, bleus, plans, etc.) peuvent être compris par certaines personnes moins facilement qu'on ne l'aurait d'abord cru. Ces représentations graphiques peuvent demander des aptitudes de traitement de l'information que certains individus ne possèdent pas et, ce qui est pire, qu'il peut être extrêmement difficile d'acquérir.

3) Si certains individus éprouvent de la difficulté à interpréter une carte topographique, n'importe qui devrait pouvoir comprendre une "carte cinématographique", qui consiste en une visite simulée d'une ville, commandée par l'utilisateur, au moyen de vidéodisques et de graphiques informatisés. Même si la carte cinématographique est une impressionnante démonstration de l'application des communications picturales à la recherche documentaire (qui présente d'énormes avantages pour les personnes confinées au logis, par exemple), il y a lieu de se demander si cette carte pourrait vraiment remplacer le plan de ville auquel nous somme habitués. Le plan de ville et la carte cinématographique font appel, semble-t-il, à des modes de traitement de l'information différents et chacun répond à des besoins différents. L'idéal serait de combiner dans un même système la carte cinématographique et la carte topographique. Cependant, on peut se demander si pareil système sera à court terme accessible à l'ensemble de la population ou s'il ne s'adressera pas plutôt à une classe hautement spécialisée.

4) Des schémas graphiques simples comme des cases, des flèches et des cercles peuvent faciliter l'appréhension de notions complexes en science, en éducation et en technique. On a émis l'hypothèse que l'utilité de schémas graphiques n'est pas arbitraire et qu'elle ne réside pas dans des conventions ou des habitudes, mais qu'elle exploite le besoin qu'a l'esprit humain de rechercher une simplicité sous-jacente et de se concentrer sur des forces et des relations génériques -- de saisir les caractéristiques structurales de phénomènes dont l'occurrence n'est ordinairement pas observable par des moyens simples lorsqu'ils se produisent.

5) Même si les schémas graphiques peuvent faciliter le processus de la pensée, les images graphiques dynamiques, c'est-à-dire des images en voie de transformation, peuvent mieux encore venir en aide à l'acquisition du savoir et à la solution de problèmes. En effet, un graphique dynamique peut faciliter les opérations de "restructuration" mentale (la transformation du modèle interne initial que l'on a d'un problème) et permettre ainsi à l'individu de saisir plus rapidement la solution appropriée. Nous avons vu, à l'aide de quelques exemples, comment des symboles dynamiques produits par ordinateur sont utilisés pour l'enseignement de notions d'électricité.

6) Pour terminer, nous avons émis l'hypothèse que, même si des graphiques dynamiques peuvent faciliter la pensée, il est utile de faire la distinction entre "mouvement" et "séquence". Le cerveau peut saisir la signification d'une "séquence" en intégrant des échantillons discrets successifs d'images fixes qui précisent les caractéristiques de transformations d'événements dynamiques sans qu'il soit nécessaire de percevoir l'illusion d'un mouvement continu de ces transformations comme au cinéma et à la télévision. La distinction entre "mouvement" et "séquence" laisse supposer qu'il faudrait trouver des techniques pour fournir sous la forme d'une succession d'images fixes des renseignements sur des séquences visuelles dynamiques. Le recours à des techniques qui s'accommoderaient ainsi d'"images fixes" se traduirait par des économies de largeur de bande d'émission et d'espace mémoire.

RÉSUMÉ ET CONCLUSIONS

Au début du rapport, nous nous sommes posé les deux questions suivantes: (1) Qu'est-ce que les théories de la perception et de la cognition peuvent nous apprendre sur la façon dont l'être humain comprend et utilise les images? (2) En quoi cette recherche s'applique-t-elle à l'utilisation d'images sur Télidon? En essayant de répondre à ces questions, nous avons touché à de nombreux aspects de la communication picturale, mais le domaine est si vaste qu'il est impossible d'explorer dans un seul exposé d'autres théories et expériences pertinentes. Le lecteur trouvera donc une bibliographie à l'annexe B.

Voici, après analyse des ouvrages, nos principales constatations relativement au rôle des images sur Télidon:

(1) Bien que des images de différentes sortes pourraient avoir d'importantes répercussions sur les systèmes vidéotex de l'avenir, y compris Télidon, l'utilisation d'une image ou d'un graphique n'accroîtra pas <u>automatiquement</u> l'utilité et l'acceptabilité de Télidon. L'utilisation d'une image entraîne des dépenses pour les fournisseurs d'information, les exploitants et les utilisateurs du système. Les concepteurs de pages, notamment, doivent se méfier des cas d'espèce où l'utilisation de l'imagerie visuelle sur Télidon impose des exigences à l'utilisateur mais n'apporte rien ou, pire encore, nuit à la compréhension des idées.

(2) Les fournisseurs d'information et les concepteurs de pages devraient éviter de mettre en conflit mots et images, c'est-à-dire de se compliquer la vie en essayant de trouver des moyens de <u>remplacer</u> des descriptions verbales claires par des images visuelles. L'utilisation d'images de différentes sortes sur Télidon trouvera sûrement son utilité, mais il est

peu probable qu'un langage purement iconique arrive un jour à supplanter les fonctions descriptives du langage. Le vrai défi pour les fournisseurs d'information est de comprendre comment les images peuvent servir de "fondement conceptuel" pour faciliter la compréhension verbale et comment le langage dirige l'interprétation des images dans différents contextes.

(3) Une image peut engendrer différentes propositions selon les différents contextes où elle est utilisée. La polysémie des images est un problème directement attribuable à la souplesse du système conceptuel humain et à sa dépendance au contexte. C'est pourquoi les images posent de sérieux problèmes dans les secteurs de la recherche appliquée qui intéressent Télidon, par exemple, la classification et la consultation automatique d'images et la mise au point d'"index iconiques". Les recherches portant sur la nature de la catégorisation humaine, tout spécialement sur le caractère prototypique des concepts, ont un rapport avec le problème de mise au point d'index verbaux et picturaux sur Télidon. On a cherché à démontrer que les meilleurs index picturaux seraient constitués d'éléments graphiques simples et abstraits néanmoins reconnaissables comme étant des représentations picturales d'objets et d'événements particuliers. Les symboles iconiques utilisés lors des Jeux olympiques de Montréal constituent un bon exemple de ce type d'index.

(4) En ce qui a trait aux images "iconiques" -- images qui renseignent sur les propriétés physiques des objets et des événements -- il ne devrait pas y avoir de problèmes de "lisibilité picturale" sur Télidon. Quiconque peut voir le monde visuel devrait être en mesure de le percevoir à travers une image. Toutefois, les images ne transmettent pas l'information à notre système visuel de la même façon que le

font les scènes réelles, car certaines formes d'adaptation sont nécessaires. Le fait d'être capable de "lire" une image ne signifie pas nécessairement que l'on perçoit le message précis qu'elle veut transmettre ou encore que quelqu'un d'autre l'interprétera de la même façon. Le contexte a une incidence directe sur la signification des images, même les images réalistes.

(5) Les images les plus faciles et les plus rapides à lire sur Télidon ne seront pas nécessairement celles qui comportent le plus haut degré de réalisme photographique. Certaines recherches théoriques et empiriques portent à croire que les croquis conservent la plupart des variables importantes de la structure optique. Il peut, de plus, être plus facile de reconnaître un objet représenté par un dessin ou par une caricature que par une photographie ou par un dessin au trait réaliste. Cela s'explique peut-être par le fait que l'artiste, à travers le croquis ou la caricature, réussit à donner une forme extérieure à la nature schématique ou "canonique" de nos codages internes des objets et des événements perceptuels. Ces hypothèses ont une utilité pratique pour Télidon, média qui excelle dans la production de graphiques schématiques de haute qualité. La nécessité du réalisme pictural sur Télidon est une question qui devra faire l'objet d'autres recherches où le degré optimal de réalisme photographique sera évalué en fonction de la nature du processus de consultation, des délais de transmission et de la "crédibilité" des différentes formes de représentation visuelle auprès des utilisateurs.

(6) Des graphiques schématiques simples peuvent jouer un grand rôle sur Télidon, notamment dans les domaines scientifiques et pédagogiques, où un graphique peut aider l'utilisateur à saisir des principes et des concepts abstraits qui ne sont pas

directement visibles au niveau superficiel des choses. Les images graphiques favorisent la réflexion et gagnent peut-être même en efficacité lorsqu'elles montrent des transformations car, à ce moment-là, elles nous aident à réaliser les opérations de "restructuration" indispensables à la pensée créatrice. Sur le plan de la recherche, on pourrait envisager de concevoir des techniques efficaces pour produire des graphiques dynamiques sur Télidon.* On a fait remarquer que les gens peuvent comprendre des séquences successives d'images fixes représentant des événements visuels sans ressentir le besoin de percevoir l'illusion du mouvement continu. Cette distinction entre "mouvement" et "séquence" laisse croire qu'il serait possible de mettre au point des techniques permettant de produire des séquences visuelles dynamiques sur Télidon sans consommation excessive de largeur de bande et d'espace mémoire.

(7) Les concepteurs de pages doivent garder à l'esprit que, si la fonction "iconique" des images ne pose pas de problèmes d'interprétation, il en va autrement de l'utilisation "symbolique" d'images schématiques pour représenter des idées abstraites qui, bien que facile à réaliser à l'aide des IDI, demande des connaissances spécialisées; par conséquent, on ne peut être assuré qu'elles seront facilement comprises par l'utilisateur moyen de Télidon.

(8) Des images graphiques telles les cartes et les plans feraient de bons sujets de recherche sur Télidon car elles se trouvent à mi-chemin entre les fonctions symboliques et iconique des

---

* Le Groupe de recherche sur Télidon a déjà mis au moins un protocole de communication graphique appelé Instructions de manipulation d'images (IMI) qui permet, entre autres, de produire des graphiques dynamiques.

des images et peuvent être plus difficiles à comprendre et à utiliser qu'on ne le croyait au départ. En fait, certaines expériences montrent qu'il existe des différences marquées dans l'aptitude des gens à utiliser des cartes et à les comprendre, et que ces différences sont enracinées dans leurs capacités cognitives innées.

En terminant, nous aimerions attirer l'attention sur un des avantages secondaires importants, selon l'optique behavioriste, que la recherche appliquée peut avoir, à notre avis, sur de nouveaux systèmes de télécommunications comme Télidon. En effet, non seulement les théories de la cognition, de la perception et de la communication peuvent-elles nous apprendre quelque chose sur Télidon, mais Télidon lui-même peut nous aider à découvrir de nouvelles questions théoriques et empiriques. Nous avons vu, par exemple, qu'avant de considérer comme une nécessité le réalisme pictural sur Télidon, il faudra faire d'autres recherches fondamentales sur "l'équivalence informationnelle" des différents types de représentations picturales, compte tenu d'un certain nombre de variables tels le processus de consultation, la crédibilité des images et le temps de transmission. L'étude des rapports qui existent entre ces variables pourrait nous aider non seulement à prendre des décisions pratiques quant à l'utilisation des images sur Télidon mais aussi à mieux comprendre les principes inhérents au système visuel humain. Ce qu'il faut retenir c'est que la question "quelles sont les répercussions?" peut avoir deux significations distinctes, c'est-à-dire

non seulement les répercussions des théories sur Télidon, mais aussi les répercussions de Télidon dans l'élaboration de théories sur le traitement de l'information par l'homme?[1]

---

[1]Télidon soulève des questions fondamentales non seulement en ce qui concerne les images mais aussi d'autres aspects fondamentaux du système humain de traitement de l'information. Pour ne citer qu'un exemple, le Groupe de recherche en comportement du ministère des Communications, tout en étudiant les méthodes optimales d'indexation et de structuration de la base de données, se trouve également aux prises avec des questions fondamentales de "représentation cognitive" -- soit trouver comment la structuration de nos schèmas internes peut servir à l'élaboration de meilleurs schèmes d'indexation. Voir Recherche en comportement sur Télidon 1, Ottawa, février, 1980, pour une analyse des projets et des questions abordés par le Groupe de recherche en comportement.

OUVRAGES DE RÉFÉRENCE

Arnheim, R. The Dynamics of Architectural Form. Berkeley: University of California Press, 1975.

_____. Art and Visual Perception. Berkeley and Los Angeles: University of California Press, 1964.

_____. Visual Thinking. Berkeley: University of California Press, 1979. (voir aussi La pensée visuelle. Paris, Flammarion, 1976.

Attneave, F. "Representation of Physical Space". In A. Melton and E. Martins (Eds.), Coding Processes in Human Memory. Washington. D.C.: V.W. Winstone & Sons, 1972.

Bartlett, F.C. Remembering. Cambridge: Cambridge University Press, 1932.

Bellugi, U. and Klima, E. The First Signs of Language. Cambridge: Harvard University Press, sous presse.

Bertin, J. Sémiologie graphique: les diagrammes, les réseaux, les cartes. Paris: Mouton, 1973.

Bliss, C.K. Semantography: Bliss Symbolics. Semantography Publications, 1965.

Bower, T.G. R. "Discrimination of Depth in Premotor Infants", Psychonomic Science, 1964, 1, 368.

_____. "Slant Perception and Shape Constancy in Infants", Science, 1966, 151, 832-834.

Bower, G.H.; Karlin, M.B., & Dueck, A. "Comprehension and Memory for Pictures". Memory & Cognition, 1975, Vol. 3(2), 216-220.

Bown, H.G.; O'Brien, C.D.; Sawchuck, W., and Storey, J.R. Instructions de description de l'image (IDI) pour le système vidéotex Télidon. Ministère des Communications. Note technique n° 699-F. Novembre 1979.

_____. Description générale de Télidon: Proposition canadienne concernant les systèmes vidéotex. Ministère des Communications. Note technique n° 697-F. Décembre 1978.

Bransford, J.D., & Johnson, M.R. "Considerations of some Problems of Comprehension". In W.G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Bregman, A.S. "Perception and Behavior as Compositions of Ideals". Cognitive Psychology, 1977, 9, 250-292.

_____. "Asking the 'What for' Question in Auditory Perception". In Kubovy, M. and Pomerantz, J.R. (eds.) Perceptual Organizations. Hillsdale, New Jersey: Lawrence Erlbaum, (sous presse).

Carmichael, L.; Hogan, H.P. & Walter, A.A. "An Experimental Study of the Effect of Language on the Reproduction of Visual Perceived Form". Journal of Experimental Psychology, 1932, 15, 73-86.

Chomsky, N. Aspects of the Theory of Syntax. Cambridge: MIT Press, 1965. ou Aspects de la théorie syntaxique. Paris, Ed. du Seuil, 1970.

Clowes, M.B. "Transformational Grammars and the Organization of Pictures". In Grasselli (Ed.) Automatic Interpretation and Classification of Images. New York: Academic Press, 1969.

Duncan, H.F.; Gourlay, N. & Hudson, W. A Study of Pictorial Perception among Bantu and White Primary School Children in South Africa. Johannesberg: Witwatersrand University Press, 1973.

Dwyer, F.M. A Guide for Improving Visualized Instruction. State College, Pa.: Learning Services, 1972.

_____. Strategies for Improving Visual Learning. State College, Pa.: Learning Services, 1978.

Fischler, M.A. "Machine Perception and Description of Pictorial Data". Proceedings of the International Joint Conference on Artificial Intelligence. Washington, D.C., 1969.

Firschein, O. and Fischler, M.A. "A Study in Descriptive Representation of Pictorial Data". Pattern Recognition. Vol. 4, 361-377.

Fodor, J. The Language of Thought. Cambridge: Harvard University Press, 1975.

Gibson, J.J. "A Theory of Pictorial Perception". Audio-Visual Communications Review, 1954, 1, 3-23.

_____. The Senses Considered as Perceptual Systems. Boston: Houton-Mifflin, 1966.

_____. "The Information Available in Pictures". Leonardo, 1971, 4, 27-35.

Gombrich, E.H. The Visual Image. Scientific American. September, 1972.

_____. Art and Illusionn. New York: Pantheon, 1960 ou L'art et l'illusion. Paris, Gallimard, 1971.

Goodman, N. Languages of Art: An Approach to a Theory of Symbols. Indianapolis: Bobbs-Merrill, 1968.

Goodman, N. "On J.J. Gibson's New Perspective". Leonardo, 1971, 4, 359-360.

Gopnik, A. "Did you Hear the One about the...: A Meta-Cognitive Theory of Humour." Non publié. Université McGill, 1979.

Hagen, M. & Jones, R. "Cultural Effects on Pictorial Perception: How many Words is one Picture Really Worth?" In R. Walk and H. Pick (Eds.), Perception and Experience. New York: Plenum, 1978.

Hebb, D.O. The Organization of Behavior. New York: Wiley, 1949.

Hochberg, J. & Brooks, V. "Pictorial Recognition as an Unlearned Ability", American Journal of Psychology, 1964, 75, 624-628.

Hochberg, J. The Representation of Things and People. In Gombrich, E.H. Hochberg, J. and Black, M. Art, Perception and Reality. Baltimore: Johns Hopkins University Press, 1972.

_____. In the Mind's Eye. In Haber, R.H. (Ed.). Contemporary Theory and Research in Visual Perception. New York: Holt, Rinehart and Winston, 1968.

Huggins, W.H. and Entwisle, D.R. Iconic Communication: An Annotated Bibliography. Baltimore: Johns Hopkins Press, 1974.

_____. Iconic Communications. IEEE Transactions on Education, 14, 1969, 158-163.

Jahoda, G. & McGurk, H. "Pictorial Depth Perception in Scottish and Ghanaian Children: A Critique of Some Findings with the Hudson Test". International Journal of Psychology, 1974, 9(4), 225-267.

Johnson, A.R. Dialogue and the Exploration of Context: Properties of an Adequate Interface. Presented at the Fourth Annual International Symposium of the American Society of Cybernetics. Washington, D.C. 1970.

Kennedy, J.M. A Psychology of Picture Perception. San Francisco: Jossey-Bass, 1974.

Kennedy, J.M. & Fox, N. "Pictures to See and Pictures to Touch." In D. Perkins and B. Leandar (Eds.). The Arts and Cognition. Baltimore: Johns-Hopkins Press, 1977.

Kolers, P.A. Some Formal Characteristics of Pictograms. American Scientist, 57, 1969, 348-363.

Lakoff, G. Images Schemas. Paper Communication presentée au Second Annual Meeting of the Cognitive Science Society, Yale University, New Haven, 1980.

Lippman, A. Movie-Maps: An Application of the Optical Videodisc to Computer Graphics. Document publié par le Architecture Machine Group, Massachusetts Institute of Technology, non daté.

Loftus, E.F. & Palmer, J.C. Reconstruction of Automobile Destruction: An Example of the Interaction Between Language and Memory. J.V.L.V.B., 1974, 13, 585-589.

Mead, M. and Modley, R. Communication among all People, Everywhere. Natural History, 77, 1968, 56-73.

Mills, M.I. The Effectiveness of Real-Time versus Still-Frame Television Channels for Interactive Problem-Solving. Systems Engineering Report 145. Bell-Northern Research, Ottawa, Ontario, 1972.

_____. On Fitting Cartoon-Strips to Descriptions: A Study of some Relations Between Perception and Language. Thèse de doctorat non publiée, Université McGill, 1980a.

_____. Cognition and the Comprehension of Cartoon-Strip Analogies. Communication présentée au National Symposium for Research in Art: Representation and Metaphor. University of Illinois, 1980b.

Mills, M.I. and Bregman, A.S. On Fitting Cartoons to Descriptions: Viewpoint Transformations in Event Comprehension. Non publié, 1980.

Minsky, M. A Framework for Representing Knowledge. In P.H. Winston (Ed.) The Psychology of Computer Vision. McGraw-Hill, 1975.

Modley, R. Pictograms and Graphs. Harper and Brothers, 1947.

Muter, P. The Role of Graphics in Learning on Videotex: Evidence from other Media. Technical Memorandum, Department of Communications, Ottawa, Ontario, 1980.

Narasimhan, R. On the Description, Generation, and Recognition of Classes of Pictures. In Graselli, A. (Ed.) Automatic Interpretation and Classification of Images. New York: Academic Press, 1969.

Newell, A. & Simon, H.A. Human Problem Solving. Englewood-Cliffs, New Jersey: Prentice-Hall, 1972.

Newport, E.L., and Bellugi, U. Linguistic Expression of Category Levels in a Visual-Gestural Language: A Flower is a Flower is a Flower. In Rosch, E. and Lloyd, B. (Eds.) Cognition and Categorization, New York: Lawrence Erlbaum, 1978.

Norman, D.A. & Rumelhart, D.E. Explorations in Cognition. San Francisco: W.H. Freeman & Sons, 1975.

Olson, R.K.; Pearl, M.; Mayfield, N. & Millar, D. "Sensitivity to Pictorial Shape Perspective in 5-year Old Children and Adults". Perception and Psychophysics, 1976, 20(3), 173-178.

Paivio, A. Imagery and Verbal Process. New York: Holt, Rinehart and Winston, 1971.

Perkins, D.N. "Caricature and Recognition". Studies in the Anthropology of Visual Communication. Spring 1975, 2(1).

_____. Pictures and the Real-Thing. Communication présentée à la Conference on Processing of Visible Language, Niagara-on-the Lake, Septembre 1979.

Perkins, D.N. & Hagen, M. "Convention, Context and Caricature". In M. Hagen (Ed.), The Perception of Pictures. Vol. 1: Alberti's Window: The Projective Model of Pictorial Information. New York: Academic Press, 1980.

Piaget, J. La psychologie de l'intelligence. (1947). Traduit sous le titre Psychology of Intelligence. London: Routledge and Kegan Paul, 1950.

Rigney, J.W. and Lutz, K.A. The Effects of Interactive Graphics Analogies on Recall of Concepts in Science. Technical Report, No. 79, Department of Psychology, University of Southern California, 1976.

Rock, I. An Introduction to Perception. New York: MacMillan, 1975.

Rosch, E. Cognitive Reference Points. Cognitive Psychology, 1975, 7, 532-547.

Rosch, E. & Lloyd, B. (Eds.) Cognition and Categorization. New York: Lawrence Erlbaum, 1978.

Ryan, T.A., & Schwartz, C.B. "Speed of Perception as a Function of Mode of Representation". American Journal of Psychology, 1956, 69, 60-69.

Shaw, R. and Bransford, J.D. Perceiving, Acting and Knowing: Toward an Ecological Psychology. N.J.: Lawrence Erlbaum, 1977.

Thorndyke, P.W., and Stasz, C. Individual Differences in Procedures for Knowledge Acquisition from Maps. Cognitive Psychology, 12, 137-175, 1980.

Winograd, T. Understanding Natural Language. N.Y.: Academic Press, 1972.

Wollen, K.A.; Weber, A. & Lowry, D.H. Bizareness versus Interaction of Mental Images as Determinants of Learning. Cognitive Psychology, 1972, 2, 518-523.

Woolfe, R. Videotex: The New Television/Telephone Information Services. London: Heyden & Son, 1980.

Verbrugge, R.R. The Comprehension of Analogy. Thèse de doctorat non publiée. University of Minnesota, 1974.

ANNEXE A
INTRODUCTION À LA PSYCHOLOGIE DE LA PERCEPTION PICTURALE

par

Adam Gopnik

Plutôt que d'entreprendre un relevé exhaustif de documents disparates, nous avons regroupé les bibliographies ci-dessous par thème et par sujet et décrit globalement la controverse qui en constitue le point de focalisation. Nous croyons que cette façon de procéder illustre mieux que ne pourrait le faire une simple liste bibliographique, les questions posées et les positions adoptées par les praticiens de cette discipline complexe et en pleine évolution.

A.- La théorie psychologique classique de la perception de l'image

1. Dergowski, J. "Illusion and Culture" in R. Gregory and E. Gombrich (Eds.) Illusion in art and nature. New York: Charles Scribner's Sons, 1973.
2. Gombrich, E.H. L'art et l'illusion. Paris, Gallimard, 1971.
2. " " Meditations on a Hobby Horse. New York: Phaidon, 1963.
3. Gombrich, E.H. "The What and the How: Perspective Representation and the Phenomenal World" in Logic and Art, Essays in Honour of Nelson Goodman, ed. Richard Rudner and Israel Schleffer. Indianapolis: Bobbs-Merrill, 1972.
4. " " "The Mask and the Face" in Art, Perception and Reality, Maurice Mandelbaum (Ed.). Baltimore: Johns Hopkins Press, 1972.
5. Goodman, N. Languages of Art: An Approach to the Theory of Symbols. Indianapolis: Bobbs-Merrill, 1968.
6. Hochberg, J. "Nativism and Empiricism in Perception". In Psychology in the Making. L. Postman (ed.). New York: Holt, Rinehart and Winston, 1962.
7. Silvers, Anita. "Show and Tell" in The Arts, Cognition and Basic Skills. St-Louis: CEMREL, 1978.
8. Wilson, B. and Wilson, M. "Recycling Symbols" in The Arts, Cognition and Basic Skills. St-Louis: CEMREL, 1978.

Prenons un dessin en perspective d'un cube. Pourquoi y voyons-nous un cube et non un curieux agencement de formes en deux dimensions? Si, comme c'est en fait le cas, on peut illustrer le même objet par un nombre infini d'agencements différents (par exemple, notre "cube" pourrait être une figure où le "devant" et le "côté" se trouveraient sur le même plan, coiffés d'un triangle drôlement fait) et si nous pouvions

logiquement interpréter cette figure d'un nombre infini de façons, pourquoi alors notre oeil ne voit-il qu'une seule figure -- celle d'un cube?

La théorie psychologique classique, du moins celle qui remonte au temps de l'évêque Berkeley, soutient que c'est l'expérience qui nous fait voir un cube dans ce dessin. Nous voyons ce que nous nous attendons de voir. Nous y voyons un cube parce que, selon notre expérience du monde, la plupart des angles sont des angles droits et les lignes parallèles convergent. Obéissant à une habitude perceptuelle, nous voyons l'objet que l'agencement de lignes et de formes capté par notre oeil serait normalement le plus susceptible d'illustrer.

En 1960, E.H. Gombrich a repris cet énoncé simple de la psychologie classique et en a fait une interprétation radicalement nouvelle dont on peut dire qu'elle est à la source d'une nouvelle discipline. Gombrich a introduit l'idée que toute représentation était le fruit d'une convention, d'une expérience codifiée sous une forme consacrée par la tradition artistique. Gombrich a avancé que nous produisons des images non pas en reproduisant ce que nous voyons, mais en manipulant des stéréotypes conventionnels de représentation qu'il a appelés "schémas". Gombrich ne voulait pas dire que toute représentation impliquait simplement le recyclage de symboles conventionnels; au contraire, toute l'histoire de l'art de la période qui a suivi la Renaissance décrit la démarche des artistes qui essayaient de modifier ou de "corriger" ces schémas conventionnels en vue d'atteindre une précision de plus en plus grande dans leurs représentations. Gombrich a fait remarquer que même la représentation la plus "photographique" aurait toujours recours aux formules conventionnelles de représentation qui résument nos connaissances du monde de la perception.

Tout en insistant constamment sur la force d'attraction exercée par les stéréotypes conventionnels, qui nous font représenter de façon familière les objets non familiers (de sorte que le dessin d'un objet non familier - comme un animal ou un paysage inconnus de l'artiste - a toujours été rendu, dans l'histoire de l'art, selon un schéma familier), Gombrich a insisté également sur la façon dont l'artiste pourrait corriger ou modifier ces schémas en vue de produire des représentations de plus en plus exactes. Il croit en la possibilité d'un naturalisme "objectif", bien qu'il croit aussi que seule la lente évolution de la formule "conventions et modifications", de la formule "schémas et corrections", peut aboutir à la fidélité associative. Par contre, une école importante de théoriciens composée principalement de philosophes, dont les travaux font suite à l'oeuvre de Gombrich, croit que le naturalisme est lui-même une pure invention de l'esprit, que les "schémas" ne sont que des symboles arbitraires pour représenter les objets qui nous entourent et que la question de la ressemblance des symboles avec les objets qu'ils sont censés représenter n'entre pas en ligne de compte. Un des plus éminents parmi ces théoriciens, Nelson Goodman, a soutenu dans son ouvrage réputé The Languages of Art, que les schémas conventionnels de l'art occidental jouent le même rôle que les mots d'une langue parlée et que leur rapport avec les objets qu'ils représentent est purement arbitraire et imposé. Selon Goodman, la représentation iconique dépend non pas de l'imitation, de l'illusion ou de l'information, mais de l'inculcation. Pratiquement n'importe quelle image peut représenter pratiquement n'importe quoi, c'est-à-dire qu'il existe ordinairement un système de représentation qu'on peut appliquer à une image et à un objet de sorte que l'image représente l'objet; d'ailleurs, il existe ordinairement plusieurs de ces systèmes (Goodman, idem, 1968). Ainsi, pratiquement n'importe quoi peut représenter pratiquement n'importe quoi d'autre. Goodman croit à l'emprise des conventions et de l'expérience, mais, à l'encontre de Gombrich, range dans deux champs distincts l'expérience du monde visuel et l'expérience

du monde iconique (iconosphère), parce que le monde visuel et le monde iconique sont différents et nous apportent une expérience essentiellement différente. Nous comprenons les images non pas d'après notre expérience du monde, mais d'après notre expérience d'autres images. Pour en revenir à notre exemple du dessin d'un cube, nous y voyons un cube non pas parce que le cube est la forme la plus familière d'un nombre infini de formes possibles que le dessin pourrait représenter, mais parce que nous savons que dans notre culture, on représente de cette façon la forme du cube. Nous nous trouvons face au symbole qui, dans notre système de conventions, représente un cube.

Au cours des dernières années, des ouvrages reprenant avec certaines variations la thèse de Goodman ont connu une certaine popularité (chose surprenante, les auteurs y ramènent souvent la position de Gombrich à celle de Goodman, malgré le fait qu'elles soient complètement opposées). Les philosophes affectionnent particulièrement la position de Goodman parce qu'elle ramène la discussion sur la façon dont fonctionnent les signes et les symboles, une question philosophique traditionnelle, et qu'elle évite de se demander si un dessin ressemble à ce qu'il est censé représenter, une question qui ennuie les philosophes. Récemment, Anita Silvers a repris plus à fond l'idée que les images sont des symboles. Elle croit que, même si l'on peut dire qu'une image ressemble en un certain sens à un objet qui nous entoure, nous utilisons les images non pas pour représenter des objets, mais pour les signifier, les illustrer et les exprimer. Brent et Marjorie Wilson ont suivi un raisonnement semblable. Dans l'article intitulé "Recycling Symbols" (le recyclage des symboles), les Wilson affirment que le langage iconique est composé d'un vocabulaire de symboles conventionnels et que, pour dessiner un objet qui ne nous est pas familier, nous appliquons à ce nouvel objet un symbole que nous connaissons déjà; nous représentons ainsi l'inconnu en termes connus. La position des Wilson ressemble à celle de Gombrich, sauf que ce dernier croit que l'on peut "corriger" les symboles conventionnels dans le sens d'un plus grand naturalisme,

alors que les Wilson croient, tout comme Goodman, que l'on ne peut que "recycler" les symboles conventionnels, les mettre dans un nouveau contexte ou les mélanger à un autre symbole leur permettant de représenter un nouvel objet. Ils font remarquer, par exemple, que les enfants apprennent à raconter des histoires illustrées non pas en apprenant à représenter leur expérience, mais plutôt en apprenant un certain nombre de formules conventionnelles de narration (comme les changements de décors, les vues à vol d'oiseau, les passages des "plans d'ensemble" aux "gros-plans", et autres moyens). La langue parlée sert encore une fois à faire l'analogie: Goodman et alii affirment que le fait de pouvoir coder en images des données sur le monde ne signifie pas que les images ressemblent au monde, pas plus que le fait de pouvoir coder en anglais des données sur le monde ne signifie que la langue anglaise "ressemble" au monde.

Parmi les appuis empiriques qu'a reçus cette thèse, les travaux de l'anthropologue Dergowski sont souvent cités. Dergowski soutient que ce qui est perçu comme une bonne représentation diffère d'une culture à l'autre (et d'un âge à l'autre, de sorte que les enfants et les adultes conçoivent différemment ce qui constitue une bonne image d'un éléphant, par exemple). Cette théorie suppose que l'interprétation des images est jusqu'à un certain point une aptitude acquise. Toutefois, Kennedy (voir plus loin) a mis en doute cette interprétation des recherches de Dergowski.

Sous le titre "Nativism and Empiricism in Perception" (nativisme et empirisme dans le domaine de la perception), Julian Hochberg a très bien décrit le cheminement intellectuel de cette école de pensée; nous verrons plus en détail sous une autre rubrique les recherches initiales de cet auteur.

Que pouvons-nous retenir de ces études? Les théories de Gombrich ont déjà commencé à avoir une incidence dans le domaine de l'éducation,

où se dessine une réaction contre l'idée que la copie "corrompt" l'oeil de l'enfant et où commence à s'implanter l'idée que l'apprentissage des schémas conventionnels au moyen de la copie a un rôle important à jouer dans l'acquisition d'une certaine "instruction" visuelle. Finalement, la question de savoir si le langage des images est universel ou dépendant de facteurs sociaux demande à être approfondie et peut avoir des ramifications importantes parce que tout procédé graphique veut être compris universellement.

B.- Le gestaltisme (la théorie de la forme)

1. Arnheim, Rudolph. Art and Visual Perception: A Psychology of the Creative Eve. Berkeley: University of California Press, 1958.
2. Arnheim, Rudolph. Vers une psychologie de l'art. Ed. Seghers, Paris, 1973.
3. " " "Introduction: Psychology and the Arts" in Perception and Pictorial Representation, Calvin F. Nodine (ed.), St-Louis: CEMREL, 1979.
4. Bernheimer, R. The Nature of Representation. New York: N.Y.U. Press, 1961.

Reprenons notre dessin d'un cube. Nous savons qu'il existe des figures semblables, qu'on appelle figures réversibles ou indéterminées, dans lesquelles les faces d'un cube peuvent être interverties; il se produit alors que ce qui semblait, à première vue, être le premier plan devient l'arrière-plan. L'important à retenir, c'est que nous ne pouvons voir qu'une version de la figure à la fois et jamais les deux en même temps. Les psychologues de la forme en concluent qu'il existe certaines règles fondamentales sous-jacentes qui régissent notre perception des images. Ils avancent que nous voyons un cube dans le dessin d'un cube non pas parce que c'est ce qui est le plus habituel d'y voir, mais plutôt parce que c'est ce qui est le plus simple d'y voir. Les psychologues de la forme soutiennent que la connaissance familière du monde ne nous aide pas toujours à interpréter les images; par exemple, des formes très familières comme les lettres peuvent devenir

impossibles à déchiffrer si on les enrobe de lignes et de fioritures inusitées. Notre interprétation des images dépend non pas des probabilités de l'expérience, mais des règles sous-jacentes de structuration mentale.

Le psychologue Rudolph Arnheim a approfondi cette théorie, qui lui a permis d'élaborer une psychologie de l'art importante et complexe. Les psychologues de la forme, comme Arnheim, accordent une importance cruciale au fait que nous ne percevons pas les images en enregistrant passivement l'information qu'elles véhiculent et en "interprétant" ensuite cette information. Nous <u>imposons</u> plutôt une structure au monde afin de le rendre significatif. Selon la théorie d'Arnheim, les tableaux sont parfaitement réussis lorsque leur structure intrinsèque correspond à la structure des rapports existant dans l'objet représenté. Prenons un exemple facile à comprendre. Nous savons qu'un artiste ne peut jamais littéralement "imiter" ou "reproduire" l'intensité réelle de la lumière réfléchie d'un paysage naturel. L'artiste peut rendre les <u>rapports</u> de luminosité et l'intensité <u>relative</u> de la lumière et créer ainsi un tableau d'une ressemblance frappante. La ressemblance ne dépendra pas de la fidélité avec laquelle l'artiste aura reproduit chacun des éléments, mais de la façon dont il aura rendu la similitude des patterns de rapports entre ces éléments. Prenons un autre exemple. Pourquoi certains tableaux représentant des chevaux au galop paraissent-ils sans vie, alors que d'autres communiquent très bien la "sensation" du mouvement"? Arnheim répond en disant que les bons tableaux rendent, dans les rapports structuraux de leurs formes et de leurs lignes, les qualités structurales inhérentes de dynamisme et de mouvement. Arnheim réagit face à la position de l'école classique en demandant comment ont été inventés en tout premier lieu les archétypes de représentation. Il prétend, quant à lui, que ces "symboles" conventionnels rendent parfaitement les qualités structurales essentielles des objets qu'ils représentent; ces symboles représentent justement ce qui prime pour l'esprit, le rapport structural de

l'ensemble, et ne tiennent pas compte de détails insignifiants et transitoires. Le "bonhomme-allumettes", par exemple, est une bonne représentation en soi parce qu'il rend les qualités structurales les plus importantes de l'objet qu'il entend représenter; nous acceptons cette forme de représentation d'un être humain non pas par tradition sociale, mais parce que cette forme répond à notre besoin psychologique inné de parcimonie et de simplicité dans les descriptions.

L'approche d'Arnheim est dite "non cognitive" parce qu'elle accorde plus d'importance à la structure de l'image elle-même qu'aux déductions mentales subséquentes que nous pourrions faire sur l'image. Ce qu'Arnheim affirme toutefois avec le plus de conviction, c'est la possibilité qu'ont les images de représenter n'importe quel savoir ou sentiment. Il croit que certaines sortes de savoir ne peuvent pas être représentées autrement que visuellement. Si nous voulons comprendre une synthèse chimique, par exemple, nous aurons assez des mots et des chiffres pour acquérir une certaine connaissance du sujet, mais une image pouvant représenter les réactions dynamiques de la synthèse reste vraiment le seul moyen efficace pour saisir l'opération. Arnheim croit qu'une étude des principes sous-jacents de la structuration mentale peut permettre aux artistes de produire des illustrations qui rejoignent mieux l'objectif visé et qui sont plus faciles à comprendre pour qui les voit (Arnheim en donne un exemple dans l'introduction de son ouvrage Perception and Pictorial Representation). Toutefois, il reste à savoir si l'on peut codifier les intuitions d'Arnheim en un ensemble de méthodes. Même à cette étape préliminaire, il n'empêche qu'Arnheim peut nous apprendre qu'il n'y a guère d'obstacles insurmontables, quelle que soit l'utilisation que nous voulons faire des images. Toutefois, il faudrait prendre note qu'il manque dans la théorie d'Arnheim le mécanisme ou l'hypothèse expliquant pourquoi certains "patterns" visuels ont la structure d'expression qu'il leur attribue. Il peut dire, par exemple, que le cercle incarne les qualités de "rétroaction" ou

d'"homéostasie"; en affirmant cela, il ne fait peut-être que se livrer à ce genre d'analyse impressionniste chère aux critiques d'art, si impressionnant que soit, en l'occurrence, le vocabulaire qu'il emprunte à la psychologie. Le lecteur averti constatera qu'entre les affirmations d'Arnheim et leur vérification dans les faits, il peut parfois y avoir une assez grande différence. Il lui arrive à l'occasion de frôler la banalité: selon lui, par exemple, les gens ressentent l'orgueil de l'attitude célèbre de Napoléon lorsqu'elle est imitée par un acteur (cette attitude "incarne les qualités structurales de l'orgueil") parce que la tête se dresse au-dessus du plan de l'interaction humaine, et que les personnes orgueilleuses veulent constamment nous en imposer.

L'aspect non cognitif des recherches de Arnheim, qui mettent l'accent sur l'analyse de la structure de l'image, a eu une profonde incidence sur les spécialistes de l'étude de la perception de l'image qui forment la principale école de pensée à l'heure actuelle; cette école tire son inspiration des recherches de J.J. Gibson, dont le point de convergence est le concept de l'"information".

C.- Théorie de l'information et perception de l'image

1. Gibson, J.J. "A Theory of Pictorial Perception" in Audio-Visual Communication Review, 1954, Vol. k, 3-23.
2. " " "The Information available in Pictures" in Leonardo, 1971, no. 4, 27/35.
3. Hagen, M.A. "Picture Perception: Towards a Theoretical Model". In Psychological Bulletin, 1974, 81, 471-479.
4. Kennedy, J.M. A Theory of Picture Perception. San Francisco: Josey Bass Inc., 1974.
5. " " "Icons and Information" in D. Olson (Ed.) National Society for the Study of Education Yearbook, 1976.
6. Pirenne, M.H. Optics, Painting and Photography. Cambridge, Cambridge University Press, 1970.
7. Ward, J. "A Piece of the Action" in Perception and Pictorial Representation. Calvin F. Nodine (ed.). St-Louis, CEMREL, 1979.

Alors que les psychologues précédents avaient été marqués par l'ambiguïté apparente de notre monde visuel et en avaient conclu que notre perception de ce monde devait être le fruit de déductions de l'esprit éclairé par l'expérience, le psychologue J.J. Gibson a mis l'accent, avec une lucidité qui ne manque pas de bon sens, sur la quantité d'information présente dans le monde visuel. Il a justement mis l'accent sur l'élément que la plupart des psychologues avaient négligé: on résout la plupart des ambiguïtés du monde visuel en s'y déplaçant; nous découvrons ainsi ce qui est constant dans le champ visuel, de même que l'origine de l'ambiguïté apparente que l'on retrouve dans un champ fixe et statique. Gibson a fait remarquer que la plupart des psychologues s'étaient fourvoyés en étudiant la perception visuelle uniquement dans un champ fixe et statique, alors qu'elle est caractérisée en fait par une rétroaction constante grâce à un débit continuel et incessant d'information.

Gibson a éprouvé des difficultés faciles à prévoir lorsqu'il s'est attaqué à la perception visuelle des images, qui sont, par leur nature même, statiques et fixes (nous ne pouvons nous promener dans une chambre d'hôpital peinte par Van Gogh). Comment se fait-il qu'une image nous soit intelligible? Gibson a d'abord avancé (dans son ouvrage de 1954) que l'image d'un objet devait son intelligibilité au fait qu'elle reproduisait en tous points la façon dont nous verrions la lumière distribuée sur l'objet s'il nous était lui-même donné à voir. Si l'on y réfléchit un peu, on se rend compte qu'un très petit nombre d'images répond à ce critère. Il arrive qu'une silhouette ou un dessin au trait nous soit intelligible sans offrir la moindre ressemblance avec la façon dont la lumière est distribuée de point en point sur l'objet représenté. Gibson a alors modifié sa théorie de la perception de l'image (dans son ouvrage de 1971) pour dire que l'image d'un objet reproduit l'information que révèle la distribution de la lumière sur l'objet représenté (démarcations, contours, gradients de textures, etc.).

J.M. Kennedy a repris cette idée et a tenté de lui donner toute la précision et la clarté possibles. Il a cherché à démontrer que, étant donné que l'on peut représenter l'information contenue dans le monde visuel à l'aide de contours dessinés, le langage des contours devient alors une sorte de lingua franca de la représentation. Selon Kennedy, tout individu, quels que soient ses antécédents culturels, peut interpréter le langage des contours, et on peut s'en servir pour représenter à peu près n'importe quoi de ce qui nous entoure. À l'instar de Gibson, Kennedy démontre que cette universalité de l'intelligibilité des contours provient du fait que le langage des contours reproduit l'information la plus importante sur le monde visuel, et non du fait que nous apprenons une série de conventions. Les éléments qui importent vraiment dans la vision sont les discontinuités, les démarcations, les contours, les endroits où une texture en remplace une autre, et les endroits où une forme en masque une autre. Si les lignes peuvent représenter les éléments fondamentaux, ou l'information importante, du monde visuel, elles offrent donc la possibilité de représenter tout ce qui est visible.

La position de Gibson et de Kennedy ressemble beaucoup à celle de Arnheim par l'importance qu'elle accorde à la structure réelle de l'image, dissociée des connaissances qui nous la rendent intelligible; leur position diffère essentiellement de celle d'Arnheim par leur façon de traiter le problème de l'abstraction. Arnheim considère qu'une représentation abstraite est souvent plus intelligible qu'une représentation réaliste parce qu'elle reproduit les rapports structuraux qui caractérisent dans notre esprit l'objet représenté. Gibson et Kennedy considèrent que l'abstraction peut aider à clarifier davantage l'information reproduite, comme on teint une lamelle pour faire ressortir les éléments importants au microscope, mais ils sont bien davantage convaincus qu'une bonne image devrait réussir à reproduire non pas une structure abstraite de rapports, mais l'information "fournie" par l'éclairage.

La plupart des combats livrés actuellement dans le domaine de la perception de l'image portent sur la notion de l'information. Pirenne a tenté de combiner les notions de structure abstraite avec la notion de reproduction de l'information en inventant l'expression "forme canonique", la forme conceptuelle des objets qui est présente dans nos esprits et qui guide notre perception du monde selon un processus cybernétique continu. D'autres théoriciens ont tenté d'étendre l'approche de Kennedy pour résoudre d'autres problèmes propres à la perception de l'image. Ward, par exemple, s'est attaqué au problème du mouvement "visible" dans les images. Contrairement à Arnheim, selon qui l'image peut transmettre la sensation du mouvement grâce à sa capacité de reproduire la structure abstraite du dynamisme, la théorie de Ward veut qu'une image soit réussie lorsqu'elle reproduit l'information importante qui est donnée à voir. Arnheim prétend qu'un bon tableau d'un cheval au galop s'éloignera considérablement de la représentation fidèle de son objet pour mettre l'accent sur la structure abstraite du mouvement au moyen d'un modèle intrinsèquement dynamique de formes et de lignes. Pour Ward par contre, un bon tableau d'un cheval au galop reproduira le plus clairement possible l'information réellement donnée à voir. C'est donc le rôle de l'abstraction et de la déformation dans la production de l'image qui est ici le point en litige: est-ce qu'en s'écartant considérablement de l'apparence naturelle de l'objet représenté, on peut plus facilement en reproduire les rapports structuraux essentiels ou, au contraire, l'image la plus réussie est-elle celle qui reproduit adroitement l'information qui est réellement donnée à voir?

Comme nous le verrons plus loin, ces questions revêtent une importance fondamentale pour notre compréhension de la nature et de l'efficacité des dessins au trait et des caricatures. En effet, même si nous contestons cet aspect des recherches de Kennedy, nous pouvons encore apprendre beaucoup de l'étude de ses théories. Kennedy insiste sur le fait que le contour est l'outil graphique le plus efficace jamais

inventé; ses recherches laissent entendre que la meilleure des techniques graphiques sera celle qui nous permettra de faire les dessins-contours les plus riches. Les travaux de Kennedy sembleraient indiquer que les techniques apparentées au croquis, qui permettent de dessiner des contours réguliers et continus, seront beaucoup plus efficaces que les méthodes de dessin par remplissage de blocs de couleurs (et nous avons, de fait, remarqué personnellement que les sujets qui ont recours aux techniques de remplissage par blocs le font presque toujours pour arriver à dessiner des contours).

Dans ses travaux en cours, Kennedy cherche à établir dans quelle mesure les aveugles sont capables d'interpréter des images. Il croit que les aveugles peuvent comprendre des représentations faites au dessin-contour en utilisant un mode tactile, ce qui laisse supposer que toute l'information que nous recevons du monde sensoriel peut être codée de telle façon qu'elle devient accessible au langage des contours, c'est-à-dire aux systèmes de représentation qui mettent l'accent sur les contours et les discontinuités, même quand ce langage est tactile plutôt que visuel. Il est permis de penser que les procédés graphiques utilisés à la télévision pourraient jouer un certain rôle dans l'application de ce genre de recherches aux personnes handicapées.

D.- La physiologie de la perception de l'image

1. Berlyne, D.E. "Ends and Means of Experimental Aesthetics" in Canadian Journal of Psychology. 1972, 26, 303-325.
2. " " Studies in the New Experimental Aesthetics. New York: Wiley, 1974.
3. Gopnik, A. and Mills, M.I., Interpreting Single Pictures as Dynamic Events. Non publié. 1980.
4. Hochberg, Julian. Art and Perception. New York: Academic Press, 1979.
5. " " "Visual Art and the Structures of the Mind" in The Arts, Cognition and Basic Skills. Stanley Majeda, ed. St-Louis: CEMREL, 1979.

6. Hochberg, Julian. "The Representation of Things and People" in Art, Perception and Reality. Maurice Mandelbaum (ed.). Baltimore: Johns Hopkins Press, 1970.
7. " " "Some of the Things that Pictures are" in C.F. Nodine (ed.). Perception and Pictorial Representation. St-Louis: CEMREL, 1979.

En même temps qu'a augmenté le nombre de travaux tentant d'analyser le genre d'information contenue dans une image, on a aussi mené des recherches récemment en vue de démontrer comment la physiologie de l'oeil et du cerveau pouvait avoir une incidence sur la perception de l'image. Les travaux de Julian Hochberg se situent au delà de toute catégorisation; cet auteur démontre, ce qui est peut-être le plus intéressant dans son oeuvre, comment des contraintes relativement simples exercées sur l'oeil et le cerveau peuvent avoir une profonde incidence sur notre interprétation des images. Dans son ouvrage "Some of the Things that Pictures Are" (Une image, c'est aussi cela), par exemple, Hochberg tente de montrer comment les artistes exploitent la différence physiologique entre la vision fovéale et la vision périphérique de l'oeil. En effet, l'oeil enregistre les détails seulement dans la très petite région de la fovea située au centre du champ visuel, alors qu'il n'enregistre que des formes imprécises et vagues dans la région périphérique de l'oeil. Dans une peinture impressionniste, selon Hochberg, le peintre crée l'effet d'un premier regard ou d'une vision nouvelle en reproduisant une scène telle que vue par la région périphérique de l'oeil, là où sont enregistrées en fait nos premières impressions à la vue d'une scène nouvelle. Hochberg exlique de la même façon comment les caricatures et les dessins-contours simplifiés atteignent leur objectif: étant donné que l'oeil peut reconnaître ces formes dans le champ de la vision périphérique, il a moins d'efforts à fournir que lorsqu'il examine une reproduction détaillée dans laquelle la fovea doit analyser chacun des éléments de l'image. Nous sommes pour ainsi dire "programmés" pour fonctionner le plus efficacement avec des formes non déterminées et non détaillées. Les travaux de Hochberg sur les conséquences graphiques de ces deux

modes de vision pourraient bien avoir une profonde incidence sur la manière optimale de composer les images et les textes pour en tirer le maximum de perceptibilité visuelle.

Comme nous le verrons dans notre analyse des ouvrages de psychologie appliquée, une représentation n'est pas réussie si elle n'exerce pas un attrait esthétique. Berlyne a tenté d'expliquer nos préférences esthétiques pour certains patterns en parlant d'un système d'"excitation corticale", qui nous ferait réagir davantage à un modèle visuel en accord avec sa complexité et sa nouveauté. Si, toutefois, le modèle est trop nouveau ou trop complexe, la réaction diminue. Du point de vue esthétique, les meilleurs modèles sont ceux dont la nouveauté est compensée par la familiarité. Il n'y a rien là de bien nouveau pour les philosophes de l'esthétique puisque c'est en fait l'idée qui sous-tend les plus anciens "canons" de la beauté; mais elle est originale parce qu'elle tente d'expliquer ces préférences en termes de physiologie, par la fatigue et l'accoutumance au stimulus.

E.- Application des images à l'enseignement

1. Dwyer, F.M. A Guide for Improving Visualized Instruction. State College, Pennsylvania, Learning Services, 1972.

2. " " "Improving Visuals for Televised Instruction" in Improving College and University Teaching, 18: 288-291.

3. Klahr, D. (Ed.) Cognition and Instruction. Hillsdale: N.J.: Lawrence Erlbaum Associates, 1976.

4. Perkins, D. and Leondar, F. The Arts and Cognition. Baltimore: Johns Hopkins Press, 1977.

5. Perkins, D. "Talk about Art" in S. Majeda (ed.) Art and Aesthetics: An Agenda for the Future. St-Louis, Missouri. CEMREL INC., 1977.

6. " " "A Definition of Caricature and Caricature and Recognition" in Studies in the Anthropology of Visual Communication, 1975, II, 1.

7. Ryan, T.A. and Schwartz, C. "Speed of Perception as a Function of the Mode of Representation". American Journal of Psychology, 1956, 69, 60-69.

À la fin des années cinquante, à cause du grand intérêt suscité dans le domaine de l'enseignement par le matériel audio-visuel, les chercheurs ont entrepris des études visant à établir si les images aidaient réellement à l'apprentissage et comment, le cas échéant, elles pourraient servir le plus efficacement.

L'étude de Ryan et Schwartz (1956) concluait que les dessins éloignés de la fidélité photographique avaient généralement une plus grande efficacité pédagogique. Ryan et Schwartz ont représenté suivant différents procédés, les mêmes objets (photographies en blanc et noir, dessins modelés au lavis, dessins-contours et bandes dessinées). Dans cette expérience, la différence de style entre les dessins-contours et les bandes dessinées est importante. Les dessins-contours essayaient, malgré l'élimination d'un grand nombre de détails et, bien entendu, de tout procédé visant à représenter l'objet sous son éclairage réel, d'en reproduire néanmoins les contours réels. Dans les bandes dessinées, par contre, on observait les conventions propres à ce genre déterminé: contours réguliers, angles exagérément arrondis et autres caractéristiques popularisées par Disney et ses disciples, faussant même, comme l'admet la tradition, les proportions relatives de l'objet représenté.

Ryan et Schwartz ont découvert que les bandes dessinées avaient besoin d'être montrés moins longtemps que toutes les autres sortes de dessins pour être identifiés par les sujets. Les photographies et les dessins au trait extrêmement précis devaient être montrés plus longtemps que toutes les autres représentations. Il est évident que certaines déformations, dont celles utilisées dans les caricatures et les bandes dessinées, peuvent faciliter la perception davantage que le respect strict de la réalité.

Francis Dwyer a poursuivi des expériences dans le même esprit pendant les quinze années qui ont suivi. Il a découvert que le fait de présenter des photographies d'organes lors du cours de médecine

n'apportait rien à la compréhension des élèves, alors que la présentation de dessins-contours des mêmes organes y contribuait de façon marquée. Les recherches de Dwyer n'ont pas porté sur les bandes dessinées, dans le sens où Disney ou Ryan et Schwartz les ont utilisés. Il est bien possible que les cartoons auraient été encore plus efficaces.

Dwyer a découvert également que l'utilisation de la couleur, même lorsqu'elle ne semblait ajouter aucune information particulière, facilitait sensiblement l'apprentissage, ce qui laisse penser que les préoccupations esthétiques peuvent avoir un rôle très important à jouer, même dans l'utilisation qui semble la plus "prosaïque" des images. Une belle image est beaucoup plus utile qu'une laide. (Nous avons remarqué, dans les travaux de Berlyne, une tentative d'explication théorique de ce phénomène pédagogique.) Dwyer s'est également rendu compte que les gens croient que les dessins très réalistes et détaillés sont les outils pédagogiques les plus efficaces, bien qu'ils soient les moins efficaces en réalité. Cette découverte porte à croire qu'il existe dans les milieux de l'enseignement une résistance fondamentale et profonde à l'utilisation des représentations graphiques "schématiques", une résistance que tout nouveau procédé graphique devra surmonter.

Dwyer s'est attaqué aux problèmes particuliers posés par la télévision dans le contexte de l'enseignement. Il a découvert qu'une fois de plus les dessins au trait étaient extrêmement efficaces (là non plus, les expériences ne portaient pas sur les cartoons dans le sens où on les connaît) et que l'abondance de détails nuisait à l'efficacité des dessins présentés à la télévision. Il s'est aperçu en même temps que, si l'on augmente les dimensions de l'image télévisée, on en diminue l'efficacité comme outil pédagogique. Il est possible que la distinction faite par Hochberg entre la vision fovéale et la vision périphérique (voir ci-dessus) puisse aider à expliquer le fait.

Plus récemment, David Perkins a porté un regard nouveau à la question des bandes dessinées et des caricatures et à leur extraordinaire efficacité mnémotechnique. Perkins a fait calquer aux sujets de son expérience des photographies de gens bien connus, surtout des hommes politiques; ces dessins étaient pour la plupart méconnaissables. Il leur a ensuite demandé d'exagérer les traits particuliers des personnages; les dessins de cette deuxième série étaient beaucoup plus reconnaissables que les premiers, beaucoup plus drôles aussi.

Comme nous pouvons le constater, ces expériences infirment toute théorie de la perception de l'image voulant que la fidélité d'une image à reproduire intégralement l'information révélée par l'éclairage de l'objet représenté soit un facteur prépondérant de l'efficacité de cette représentation. Kennedy a laissé entendre que le secret de l'efficacité des bandes dessinées et des caricatures réside simplement dans l'exagération du langage des dessins au trait. Toutefois, une étude approfondie pourrait bien nous apprendre qu'il existe d'autres facteurs que le simple agrandissement géométrique de détails importants. Selon Hochberg, les caricatures seraient efficaces parce qu'elles correspondent plus étroitement à l'"image canonique" des choses, à ce souvenir visuel schématisé présent dans notre esprit. Récemment, Gopnik et Mills (non daté) ont proposé l'idée que l'efficacité des bandes dessinées dépend souvent du fait qu'elles peuvent résumer une séquence visuelle dont la perception est normalement échelonnée dans le temps (par exemple, l'expression d'un visage, l'action ou la possibilité d'action d'un personnage). Les bandes dessinées sont-elles simplement un prolongement du langage des dessins-contours, ou parlent-elles un langage différent? Voilà une question qui demande à être approfondie.

ANNEXE B

BIBLIOGRAPHIE SUPPLÉMENTAIRE

Anderson, John R. et Paulson, Rebecca. "Interference in Memory for Pictorial Information." Cognitive Psychology. Vol. 10(a), Avril 1978, 178-202.

Antes, James R. "Components of Pictorial Informativeness." Perceptual and Motor Skills. Vol. 47 (2), 459-464.

Bagget, Patricia. "Memory for Explicit and Implicit Information in Picture Stories." Journal of Verbal Learning and Verbal Behaviour. Vol. 14 (5), Octobre 1975, 538-548.

Bahrick, Harry P. et Gharrity, Kathleen. "Interaction among Pictorial Components in the Recall of Picture Captions." Journal of Experimental Psychology : Human Learning and Memory. Vol. 2, Mars 1976, 103-111.

Blake, Tyler. "Motion in Instructional Media: Some Subject-Display Mode Interactions." Perceptual and Motor Skills. Vol. 44 (3), Juin 1977, 975-985.

Bruner, J.S. et Mackworth, N.H. "How Adults and Children Search and Recognize Pictures." Human Development. Vol. 13, 1970, 149-177.

Cornell, Samuel D. "Effects on Retention of Redundancy and Simultaneity of Aural and Visual Messages." Dissertations Abstracts International. Vol. 36 (6B), Décembre 1975, 2997-2998.

Clark, Robert C. "An Experimental Study of the Effects of Five Audio-Visual Presentations Upon Mental Imagery and Learning." D.A.I. Vol. 37, Décembre 1976, 3044-3045.

Dallenbach, Robert B. "A Study of Verbal and Imaginal Memory of Children and Adults Using Human Faces, Objects and Words in Paired Associate Recognition Tasks." D.A.I., 1974, 3985.

Ferguson, Eugene S. "The Mind's Eye: Nonverbal Thought in Technology." Science. Vol. 197, Août 1977, 827-836.

Fleming, Malcolm C. "The Picture in your Mind." A.V. Communications Review. Vol. 25, 1977, 43-62.

Franken, R.E. et Davis, J. "Predicting Memory for Pictures from Rankings of Interestingness, Pleasingness, Complexity, Figure/Ground and Clarity." Perceptual and Motor Skills. Vol. 41, Août 1975.

Franken, R.E. et Rowland, G.L. "Organizational Factors in Picture Recognition Memory." Acta Psychologica. Vol. 42, Novembre 1978, 441-452.

Fredrico, E.A. et Montague, W. "Recognition Memory as a Function of Encoding Strategy and Stimulus Codability." Journal of Experimental Psychology. Vol. 104, Novembre 1975, 660-668.

Friedman, S.L. et Stevenson, M. "Perception of Movement in Pictures." Dans "What Then, are Pictures?: The Psychology of Representational Art. M.A. Hagen (éd.) (présentement sous presse).

Hagen, M.A. et Elliot, H.B. "An Investigation of the Relationship Between Viewing Condition and Preference for True and Modified Linear Perspective." Journal of Experimental Psychology: Human Perception and Performance. 1976, 479-490.

Honeck, R.P. "Five Questions on Semantics." Bulletin de Psychologie. 1976, 116-124.

Hunter, Patricia A. "Mental Imagery and Learning in a Community College: A Study to Determine if the Ability to Image Mentally Affects Learning from Pictorial or Verbal Presentations." D.A.I. Vol. 37, Décembre 1976, 3031-3032.

Just, Marcel A. et Carpenter, Patricia A. "The Semantics of Locative Information in Pictures and Mental Images." British Journal of Psychology. Vol. 66, Novembre 1975, 427-441.

Kosslyn, Stephen M. et Locklear, Eileen D. "Childrens Drawings as Data about Internal Representations." Journal of Experimental Child Psychology. Vol. 23, Avril 1977, 191-211.

Kolers, P.A. "Some Psychological Aspects of Pattern Recognition." Dans Recognizing Patterns, P.A. Kolers et M. Eden (éd.) Cambridge: M.I.T. Press, 1968.

Kolers, P.A. "Reading Pictures and Reading Texts." Dans The Arts and Cognition. Baltimore: Johns Hopkins Press, 1977.

Kriska, Steven. "Representing Multivariate Data Points as Schematic Faces." D.A.I. Vol. 37, Novembre 1976, 2545-2546.

Long, Robert E. "The Effects of Imagery Stimulation in the Motor Performance of Blind and Sighted Subjects." Dissertations Abstracts International. Vol. 35, Décembre 1974, 3554-3555.

Marks, L.E. "On Coloured-Hearing Synesthesia: Cross-Modal Translations of Sensory Dimensions." Psychological Bulletin. Vol. 82, Mai 1975, 303-331.

Mandler, Jean M. et Johnson, Nancy S. "Some of the Thousand Words a Picture is Worth." Journal of Experimental Psychology. Septembre 1976, 529-540.

Moles, Abraham A. "Image and Visual Communication - Towards a Methodology." Communications. Vol. 2, 1976, 330-348.

Nelson, Douglas R.; Reed, Valerie S. et Walling, John R. "Pictorial Superiority Effect." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2, Septembre 1976, 523-528.

Nelson, Thomas M. "Enjoyment of Perspective Transformations Induced by Movement." Canadian Journal of Behavioural Sciences. Vol. 8, Juillet 1976, 287-297.

Neisser, U. et Kerr, N. "Spatial and Mnemonic Properties of Visual Images." Cognitive Psychology. Vol. 5, 1973, 138-150.

Paivio, Allan. "Perceptual Comparisons Through the Mind's Eye." Memory and Cognition. Vol. 3, Novembre 1975, 635-647.

Paivio, Allan. Imagery and Verbal Processes. Toronto: Hold, Rinehart and Winston, 1971.

Pellegrino, J.W.; Siegel, A.W. et Dhawan, M. "Differential Distraction Effects Through Short Term and Long Term Retention of Pictures and Words." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2 (5), Septembre 1976, 541-547.

Pickford, R.W. Psychology and Visual Aesthetics. London: Hutchison Educational Service. 1972.

Pittenger, John B. et Shaw, Robert E. "Aging Faces as Viscal Elastic Events: Implications for a Theory of Non-Rigid Shape Perception." Journal of Experimental Psychology: Human Perception and Performance. Vol. 104, Novembre 1975, 374-382.

Potter, Mary C. "Short Term Conceptual Memory for Pictures." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2, Septembre 1976, 509-522.

Potter, Mary C. et Faulconer, B.A. "Time to Understand Pictures and Words." Nature. Vol. 253, Février 1975, 437-438.

Rigney, Joseph W. et Lutz, Kathy A. "Effect of Graphic Analogies of Concepts in Chemistry on Learning and Attitudes." Journal of Educational Psychology. Vol. 68, Juin 1976, 305-311.

Scull, L. Robert. "Instructional Effectiveness of Two Types of Visual Illustration in Facilitating Specific Geographic Learning at the College Level." Dissertations Abstracts International. Vol. 35, Mai 1975, 7038-7039.

Shaver, Phillip; Pierson, Lee et Lang, Stephen. "Converging Evidence for the Functional Significance of Imagery in Problem Solving." Cognition: International Journal of Cognitive Psychology. Vol. 3, 1974-75, 359-375.

Shepherd, R.N. et Podgorny, P. "Cognitive Processes that Resemble Perceptual Processes." Dans Handbook of Learning and Cognitive Processes. W.K. Estes, (éd.) Hillsdale, N.J.: Erlbaum Associates, 1977.

Shepherd, R.N. et Chipman, S. "Second Order Isomorphism of Internal Representations: Shapes of States." Cognitive Psychology. 1970, 1-11.

Shumway, Helen. "An Experimental Study in the Teaching of Visual Structural Unity by the Use of Film." Dissertation Abstracts International. Vol. 35, Mars 1975, 5791-5792.

Standing, L. et Smith, P. "Verbal-Pictorial Transformations in Recognition Memory." Canadian Journal of Psychology. Vol. 29, Décembre 1975, 316-326.

Snodgrass, Joan G. et Assiaghi, Anthony. "The Pictorial Superiority Effect in Recognition Memory." Bulletin of the Psychonomic Society. Vol. 100, Juillet 1977, 1-4.

Sullivan, A.M. "The Relative Effectiveness of Instructional Television." Interchange 1976-77. Vol. 7, 46-51.

Vitz, Paul C. "Visual Science and Modernist Art." Dans Perception and Pictorial Representation. Calvin F. Nodine (éd.) New York: Praeger, 1979.

Walk, Richard D. "La perception de l'image: une intégration empirique." Bulletin de Psychologie. Vol. 30, Mai-Juin 1977, 737-738.

Yunurmey, A. Daniel. "Hypermesia for Pictures but nor for Concrete or Abstract Words". Bulletin of Psychonomic Society. Vol. 8. Août 1976, 115-117.

Canada